

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
4 November 2004 (04.11.2004)

PCT

(10) International Publication Number  
WO 2004/094667 A2

(51) International Patent Classification<sup>7</sup>: C12Q 1/68 (74) Agent: DREW & NAPIER LLC; 20 Raffles Place, #17-00 Ocean Towers, Singapore 048620 (SG).

(21) International Application Number:

PCT/SG2004/000103

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 21 April 2004 (21.04.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/464,115	21 April 2003 (21.04.2003)	US
60/464,345	22 April 2003 (22.04.2003)	US
60/464,643	23 April 2003 (23.04.2003)	US
60/464,965	24 April 2003 (24.04.2003)	US
60/496,016	19 August 2003 (19.08.2003)	US

(71) Applicant (for all designated States except US): GENOME INSTITUTE OF SINGAPORE [SG/SG]; 60 Biopolis Street, Genome #02-01, Singapore 138672 (SG).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): REN, Ee Chee [MY/SG]; 62 Greenleaf View, Singapore 279301 (SG). HUBBERD, Martin, L. [GB/SG]; 1 Holland Avenue #04-05, Willynvile, Singapore 278965 (SG). NG, Lisa Fong Poh [SG/SG]; Blk 10B #22-06 Braddell View, Singapore 579721 (SG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/094667 A2

(54) Title: REAGENTS AND METHODS FOR DETECTING SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS

(57) Abstract: This invention provides reagents and methods for detecting the severe acute respiratory syndrome coronavirus (SARS CoV). This invention also provides related compositions, kits, systems, and computers.

## REAGENTS AND METHODS FOR DETECTING SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS

### CROSS-REFERENCES TO RELATED APPLICATIONS

5 [0001] This application claims the benefit of U.S. Provisional Application Nos. 60/464,115, filed April 21, 2003; 60/464,345, filed April 22, 2003; 60/464,643, filed April 23, 2003; 60/464,965, filed April 24, 2003; and 60/496,016, filed August 19, 2003, the disclosures of which are each incorporated by reference in their entirety for all purposes.

### FIELD OF THE INVENTION

10 [0002] The present invention relates to the fields of molecular biology and nucleic acid chemistry. The invention provides methods and reagents for detecting the severe acute respiratory syndrome coronavirus (SARS CoV) and accordingly, also relates to the fields of medical diagnostics and prognostics.

### BACKGROUND OF THE INVENTION

15 [0003] Severe acute respiratory syndrome (SARS) was first identified in late November 2002 in Guangdong Province, China. In the ensuing months, major outbreaks were reported in other parts of China, Vietnam, Canada, Singapore, Taiwan, and elsewhere in the world. The disease is unusual in its high level of infectivity, as 20 demonstrated among the health care workers and family members that have been in close contact with infected individuals. In addition, it has also been reported that infected patients do not respond to empirical antimicrobial treatment for acute community-acquired typical or atypical pneumonia (Peiris et al. (2003) "Coronavirus as a possible cause of severe acute respiratory syndrome," Lancet 361:1319-1325, which 25 is incorporated by reference).

[0004] The cause of SARS has been identified as a novel coronavirus (CoV) (Drosten et al. (2003) "Identification of a novel coronavirus in patients with severe acute respiratory syndrome," N. Engl. J. Med. 348:1967-1976, which is incorporated by reference), because clinical specimens from patients infected with SARS revealed 30 the presence of crownshaped CoV particles. This new CoV has thus been referred to as

4/10/05  
JC20 Rec'd PCT/PTO 07 OCT 2005

SARS CoV. The full-length genome sequence of the SARS CoV has been reported from different isolates, and the genome organization of SARS CoV was found to be similar to that of other CoVs (Marra et al. (2003) "The genome sequence of the SARS-associated coronavirus," Science 300:1399–1404 and Ruan et al. (2003) "Comparative 5 full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection," Lancet 361:1779–1785, which are both incorporated by reference).

[0005] CoVs are a family of positive-strand RNA-enveloped viruses called *Coronaviridae*, which are now categorized under the newly established order 10 *Nidovirales*. This order comprises the families *Coronaviridae* and *Arteriviridae*. The name *Nidovirales* comes from the Latin word *nidus*, for nest, referring to the 3'-coterminal "nested" set of subgenomic mRNAs produced during viral infection (Cavanagh (2003) "Nidovirales: a new order comprising *Coronaviridae* and 15 *Arteriviridae*," Arch. Virol. 14:629–633, which is incorporated by reference). The SARS CoV genome is very large, 29.7 kb (Marra et al. (2003), *supra*, and Ruan et al. 20 (2003), *supra*, which are both incorporated by reference), and encodes 23 putative proteins. Major structural proteins include nucleocapsid, spike, membrane, and small envelope. Nonstructural proteins include the papainlike proteinase, 3C-like proteinase, RNA-dependent RNA polymerase (RdRp), helicase, and many other proteins involved 25 with viral replication and transcription (Cavanagh (2003), *supra*, and Ng et al. (2002) "Membrane association and dimerization of a cysteine-rich, 16-kilodalton polypeptide released from the C-terminal region of the coronavirus infectious bronchitis virus 1a polyprotein," J. Virol. 76:6257–6267, which are both incorporated by reference). In other CoVs, many of the nonstructural proteins are only slightly conserved in the viral sequence, the exception being RdRp, which is highly conserved in many CoVs.

[0006] The use of oligonucleotide sequences as primers and/or probes for the 30 recognition of infectious agents is one alternative to problematic immunological identification assays and other pre-existing methodologies. For example, nucleic acid probes complementary to targeted nucleic acid sequences have been used in hybridization procedures, such as Southern blots and dot blots, to detect the target nucleic acid sequence. Many of these hybridization procedures have depended on the

cultivation and/or enrichment of the organism and, thus, are generally unsuitable for rapid diagnosis. The advent of techniques for the rapid amplification of specific nucleic acid sequences, such as the polymerase chain reaction among many others, has provided a mechanism to use primer and probe nucleic acids directly on clinical 5 specimens, thereby eliminating enrichment and *in vitro* culturing of the pathogen prior to performing the hybridization assay. Thus, amplification-based hybridization assays can provide simple and rapid diagnostic techniques for the detection of pathogens, such as the SARS CoV in clinical samples.

#### SUMMARY OF THE INVENTION.

10 [0007] The present invention provides methods and reagents for the rapid, reliable, and highly sensitive detection of the severe acute respiratory syndrome coronavirus without substantial detection of, or cross-reactivity with, other species in the coronavirus genus or species from other genera. Further, the methods and reagents of the invention can be utilized to detect the virus independent of the particular SARS 15 CoV viral type. In addition to oligonucleotides, compositions, and reaction mixtures, the invention also relates to kits and systems for detecting these pathogenic agents, and to related computer and computer readable media.

[0008] In one aspect, the invention provides an oligonucleotide consisting of a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 20 and 15-24 or complements thereof. In another aspect, the invention provides an oligonucleotide comprising a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24 and complements thereof, which oligonucleotide has 100 or fewer nucleotides. In still another aspect, the invention provides an oligonucleotide that includes a nucleic acid having at least 90% sequence 25 identity (e.g., at least 95%, etc.) to one of SEQ ID NOS: 1-12 and 15-24 or a complement thereof, which oligonucleotide has 100 or fewer nucleotides. Typically, these oligonucleotides are primer nucleic acids, probe nucleic acids, or the like in these embodiments. In certain embodiments, the oligonucleotides described herein have sequences between about 12 and about 50 nucleotides in length. For example, these 30 oligonucleotides have 40 or fewer nucleotides (e.g., 35 or fewer nucleotides, 30 or fewer nucleotides, etc.) in some embodiments. To further illustrate, the

oligonucleotides comprise at least one modified nucleotide in some embodiments. In certain embodiments, for example, at least one nucleotide of an oligonucleotide described herein is modified to alter nucleic acid hybridization stability relative to unmodified nucleotides. In some embodiments, the oligonucleotides described herein 5 comprise at least one label and/or at least one quencher moiety. In certain embodiments, the oligonucleotides include at least one conservatively modified variation.

[0009] In another aspect, the invention relates to a method of detecting a severe acute respiratory syndrome coronavirus in a sample. The method includes (a) 10 contacting nucleic acids from the sample with at least one primer nucleic acid comprising at least one nucleic acid selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24 and complements thereof in at least one nucleic acid amplification reaction (e.g., a nested polymerase chain reaction or the like). In addition, the method also includes (b) detecting the nucleic acids and/or one or more 15 amplicons thereof from the nucleic acid amplification reaction during or after (a), thereby detecting the severe acute respiratory syndrome coronavirus in the sample. The sample is typically derived from a mammalian subject, such as a human subject. In certain embodiments, at least one of the amplicons is about 440 nucleotides in length. In some embodiments, at least one round of the nucleic acid amplification reaction is 20 performed using primer nucleic acids comprising sequences selected from SEQ ID NOS: 11 or 22 and SEQ ID NOS: 12 or 20. In certain embodiments, at least one round of the nucleic acid amplification reaction is performed using primer nucleic acids comprising sequences selected from SEQ ID NOS: 15 or 18 and SEQ ID NOS: 16 or 19. Optionally, at least one of the primer nucleic acids comprises a modified primer 25 nucleic acid and/or comprises at least one label. In some embodiments, for example, the modified primer nucleic acid comprises a nucleic acid amplification specificity altering modification and/or a restriction site linker modification. In certain embodiments, (b) comprises monitoring binding between the amplicons and at least one oligonucleotide having a sequence selected from the group consisting of: SEQ ID NOS: 30 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. Typically, the oligonucleotide

comprises at least one label and/or at least one quencher moiety. In certain embodiments, for example, the oligonucleotide comprises a 5'-nuclease probe having a sequence selected from SEQ ID NO: 27 or SEQ ID NO: 28. In these embodiments, (b) optionally comprises detecting a detectable signal produced by the label, or amplifying 5 a detectable signal produced by the label to produce an amplified signal and detecting the amplified signal.

[0010] In still another aspect, the invention provides a method of determining a presence of a severe acute respiratory syndrome coronavirus in a sample, which method comprises (a) contacting nucleic acids and/or amplicons thereof from the sample with 10 one or more oligonucleotides that comprise at least one nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. The method also includes (b) monitoring (e.g., at a single time point, at 15 multiple discrete time points, continuously over a selected time period, etc.) binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides; in which detectable binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides, determines the presence of the severe acute respiratory syndrome coronavirus in the sample. The presence of the severe acute respiratory syndrome 20 coronavirus in the sample is generally unknown or unsubstantiated before (a). In certain embodiments, (a) comprises contacting the nucleic acids and/or amplicons thereof with the oligonucleotides in solution at a temperature of at least 42°C for at least 15 minutes in which a total weight of the solution comprises about 50% formalin and comprises heparin at a concentration of about 1 mg/ml. Moreover, the method 25 typically comprises a reaction other than a sequencing reaction. The sample is generally derived from a mammalian subject, such as a human subject. In certain embodiments, the nucleic acids and/or amplicons thereof and the oligonucleotides are contacted in solution. Optionally, a solid support comprises the nucleic acids and/or amplicons (e.g., arrayed on the solid support). As an additional option, a solid support 30 comprises the oligonucleotides.

[0011] In some embodiments, at least one segment of the nucleic acids is amplified prior to or during (a) using at least one nucleic acid amplification technique to produce the amplicons and (b) comprises monitoring the binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides, during or after amplification.

5 For example, the nucleic acid amplification technique typically comprises a polymerase chain reaction (e.g., a nest polymerase chain reaction, etc.), a ligase chain reaction, and/or the like. In these embodiments, the segment is optionally amplified using at least one primer nucleic acid comprising a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof in which the 10 variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. In some of these embodiments, the primer nucleic acid comprises at least one label, as described herein or otherwise known in the art. Optionally, the primer nucleic acid comprises a modified primer nucleic acid (e.g., a nucleic acid amplification specificity altering 15 modification, a restriction site linker, and/or the like).

[0012] In another aspect, the invention relates to a composition comprising a sample derived from a subject and at least one oligonucleotide that comprises a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% 20 sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, which oligonucleotide consists of 100 or fewer nucleotides. A presence of a severe acute respiratory syndrome coronavirus in the sample is generally unknown or unsubstantiated. Typically, the oligonucleotides comprise at least one chemically synthesized nucleic acid. In certain embodiments, at 25 least one nucleotide of a given oligonucleotide is modified (e.g., to alter nucleic acid hybridization stability relative to unmodified nucleotides).

[0013] Typically, at least one of the oligonucleotides comprises at least one label and/or at least one quencher moiety. To illustrate, the label optionally comprises a fluorescent dye, a weakly fluorescent label, a non-fluorescent label, a colorimetric 30 label, a chemiluminescent label, a bioluminescent label, an antibody, an antigen, biotin, a hapten, a mass-modifying group, a radioisotope, an enzyme, or the like.

[0014] The oligonucleotides of the compositions of the invention are provided in various formats. In some embodiments, for example, at least one of the oligonucleotides is in solution. In other embodiments, a solid support comprises at least one of the oligonucleotides. In these embodiments, the oligonucleotides are non-covalently or covalently attached to the solid support. Exemplary solid supports utilized in these embodiments are optionally selected from, e.g., a plate, a microwell plate, a bead, a microbead (e.g., a magnetic microbead, etc), a tube (e.g., a microtube, etc.), a fiber, a whisker, a comb, a hybridization chip, a membrane, a single crystal, a ceramic layer, a self-assembling monolayer, and the like.

5

10 [0015] To further illustrate, the oligonucleotides are optionally conjugated with biotin or a biotin derivative and the solid support is optionally conjugated with avidin or an avidin derivative, or streptavidin or a streptavidin derivative. In some embodiments, a linker attaches the oligonucleotides to the solid support. The linker is typically selected from, e.g., an oligopeptide, an oligonucleotide, an oligopolyamide, an oligoethyleneglycerol, an oligoacrylamide, an alkyl chain, and the like. Optionally, a cleavable attachment attaches the oligonucleotides to the solid support. The cleavable attachment is generally cleavable by, e.g., heat, an enzyme, a chemical agent, electromagnetic radiation, etc.

15

20 [0016] In another aspect, the invention provides a kit that includes (a) at least one oligonucleotide that comprises a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, which oligonucleotide consists of 100 or fewer nucleotides; and one or more of: (b)

25 instructions for determining a presence of a severe acute respiratory syndrome coronavirus in a sample by monitoring binding between nucleic acids and/or amplicons thereof from the sample and the oligonucleotide in which the presence of a severe acute respiratory syndrome coronavirus in the sample is unknown or unsubstantiated, or (c) at least one container for packaging at least the oligonucleotide. In some embodiments,

30 the kit further includes at least one enzyme (e.g., a polymerase, etc.) and/or one or more nucleotides (e.g., deoxyribonucleotides, etc.).

[0017] In some embodiments, the oligonucleotide is in solution, whereas in others, a solid support comprises the oligonucleotide. The solid support is optionally selected from, e.g., a plate, a microwell plate, a bead, a microbead, a tube, a fiber, a whisker, a comb, a hybridization chip, a membrane, a single crystal, a ceramic layer, a self-assembling monolayer, or the like.

[0018] Typically, the oligonucleotide comprises at least one label and/or at least one quencher moiety. Exemplary labels include, e.g., a fluorescent dye, a weakly fluorescent label, a non-fluorescent label, a colorimetric label, a chemiluminescent label, a bioluminescent label, an antibody, an antigen, biotin, a hapten, a mass-modifying group, a radioisotope, an enzyme, or the like.

[0019] In still other aspects, the invention provides a system (e.g., an automated system) for detecting a severe acute respiratory syndrome coronavirus in a sample. The system includes (a) at least one oligonucleotide that comprises a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, which oligonucleotide consists of 100 or fewer nucleotides. The system also includes (b) at least one detector that detects binding between nucleic acids and/or amplicons thereof from the sample and the oligonucleotide, and (c) at least one controller operably connected to the detector, which controller comprises one or more instruction sets that correlate the binding detected by the detector with a presence of the severe acute respiratory syndrome coronavirus in the sample. In addition, the oligonucleotide typically comprises at least one label and/or at least one quencher moiety. In some embodiments, at least one container or solid support comprises the oligonucleotide. In these embodiments, the system optionally further includes (d) at least one thermal modulator operably connected to the container or solid support to modulate temperature in the container or on the solid support, and/or (e) at least one fluid transfer component that transfers fluid to and/or from the container or solid support, e.g., for performing one or more nucleic acid amplification techniques in the container or on the solid support, etc.

[0020] In yet another aspect, the invention provides a system that includes (a) computer or computer readable medium comprising a data set that comprises a plurality of character strings that correspond to a plurality of sequences that correspond to one or more of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof  
5 wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. The system also includes (b) an automatic synthesizer coupled to an output of the computer or computer readable medium, which automatic synthesizer accepts instructions from the computer or computer readable medium, which instructions direct synthesis of one or  
10 more nucleic acids that correspond to one or more character strings in the data set.

## APPENDIX

[0021] This application is being filed with a paper appendix totaling 22 pages. This appendix provides additional description of certain aspects of the invention and is incorporated by reference.

## 15 BRIEF DESCRIPTION OF THE DRAWINGS

[0022] Figure 1 shows a nucleotide sequence of the SARS-CoV genome as determined by the Centers for Disease Control (Atlanta, GA, USA).

[0023] Figure 2 shows alignments of nucleotide sequences from various SARS-CoV isolates and from certain non-target organisms relative to one another.

20 [0024] Figure 3 schematically depicts a nested PCR procedure for detecting the SARS-CoV in a sample.

[0025] Figure 4 is a block diagram showing a representative example system for detecting the SARS-CoV in a sample.

25 [0026] Figure 5 is a block diagram showing a representative example system including a computer and a computer readable medium in which various aspects of the present invention may be embodied.

[0027] Figure 6 is a photograph of an agarose gel that shows the results of a nested polymerase chain reaction in a positive clinical sample according to the procedure of Example 2 using primers of SEQ ID NO: 11 and SEQ ID NO: 12 as

reagents 6 and 7 and using primers of SEQ ID NO: 13 and SEQ ID NO: 14 as reagents 8 and 9, respectively.

## DETAILED DESCRIPTION

### I. DEFINITIONS

5 [0028] Before describing the present invention in detail, it is to be understood that this invention is not limited to particular oligonucleotides (e.g., primer nucleic acids, probe nucleic acids, etc.), methods, compositions, reaction mixtures, kits, systems, computers, or computer readable media, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing 10 particular embodiments only, and is not intended to be limiting. Further, unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. In describing and/or claiming the present invention, the following terminology and grammatical variants will be used in accordance with the definitions 15 set forth below.

20 [0029] A "5'-nuclease probe" refers to an oligonucleotide that comprises at least two labels and emits radiation of increased intensity after one of the two labels is cleaved or otherwise separated from the probe. In certain embodiments, for example, a 5'-nuclease probe is labeled with two different fluorescent dyes, e.g., a 5' terminus reporter dye and the 3' terminus quenching dye or moiety. When the probe is intact, energy transfer typically occurs between the two fluorophores such that fluorescent emission from the reporter dye is quenched. During an extension step of a polymerase chain reaction, for example, a 5'-nuclease probe bound to a template nucleic acid is 25 cleaved by the 5' nuclease activity of, e.g., a Taq polymerase such that the fluorescent emission of the reporter dye is no longer quenched. Exemplary 5'-nuclease probes are described in, e.g., U.S. Pat. No. 5,210,015, entitled "HOMOGENEOUS ASSAY SYSTEM USING THE NUCLEASE ACTIVITY OF A NUCLEIC ACID POLYMERASE," issued May 11, 1993 to Gelfand et al., U.S. Pat. No. 5,994,056, entitled "HOMOGENEOUS METHODS FOR NUCLEIC ACID AMPLIFICATION 30 AND DETECTION," issued November 30, 1999 to Higuchi, and U.S. Pat. No. 6,171,785, entitled "METHODS AND DEVICES FOR HOMOGENEOUS NUCLEIC

ACID AMPLIFICATION AND DETECTOR," issued January 9, 2001 to Higuchi, which are each incorporated by reference.

[0030] The term "alteration" refers to a change in a nucleic acid sequence, including, but not limited to, a substitution, an insertion, and/or a deletion.

5 [0031] An "amplification reaction" refers to a primer initiated replication of one or more target nucleic acid sequences or complements thereto.

[0032] An "amplicon" refers to a molecule made by copying or transcribing another molecule, e.g., as occurs in transcription, cloning, and/or in a polymerase chain reaction ("PCR") (e.g., nested PCR, strand displacement PCR amplification (SDA),

10 duplex PCR amplification, etc.) or another nucleic acid amplification technique.

Typically, an amplicon is a copy of a selected nucleic acid (e.g., a template or target nucleic acid) or is complementary thereto.

[0033] An "amplified signal" refers to increased detectable signal that can be produced in the absence of, or in conjunction with, an amplification reaction.

15 Exemplary signal amplification techniques are described in, e.g., Cao et al. (1995) "Clinical evaluation of branched DNA signal amplification for quantifying HIV type 1 in human plasma," AIDS Res Hum Retroviruses 11(3):353-361, and in U.S Pat. No. 5,437,977 to Segev, U.S. Pat. No. 6,033,853 to Delair et al., and U.S. Pat. No. 6,180,777 to Horn, which are each incorporated by reference.

20 [0034] "Antibody" refers to a polypeptide substantially encoded by at least one immunoglobulin gene or fragments of at least one immunoglobulin gene, that can participate in detectable binding with a ligand. The term includes naturally-occurring forms, as well as fragments and derivatives. Fragments within the scope of the term as used herein include those produced by digestion with various peptidases, such as Fab, 25 Fab' and F(ab')<sup>2</sup> fragments, those produced by chemical dissociation, by chemical cleavage, so long as the fragment remains capable of detectable binding to a target molecule, such as an antigen indicative of a disease.

[0035] An "array" refers to an assemblage of elements. The assemblage can be spatially ordered (a "patterned array") or disordered (a "randomly patterned" array). 30 The array can form or comprise one or more functional elements (e.g., a probe region on a microarray) or it can be non-functional.

[0036] The term "attached" or "conjugated" refers to interactions and/or states in which material or compounds are connected or otherwise joined with one another.

These interactions and/or states are typically produced by, e.g., covalent bonding, ionic bonding, chemisorption, physisorption, and combinations thereof. In certain

5 embodiments, for example, oligonucleotides are attached to solid supports. In some of these embodiments, an oligonucleotide is conjugated with biotin (i.e., is biotinylated) and a solid support is conjugated with avidin such that the oligonucleotide attaches to the solid support via the binding interaction of, e.g., biotin and avidin.

[0037] Molecular species "bind" when they associate with one another via

10 covalent and/or non-covalent interactions. For example, two complementary single-stranded nucleic acids can hybridize with one another to form a nucleic acid with at least one double-stranded region. To further illustrate, antibodies and corresponding antigens can also non-covalently associate with one another.

[0038] The term "cleavage" refers to a process of releasing a material or

15 compound from attachment to another material or compound. In certain embodiments, for example, oligonucleotides are cleaved from, e.g., a solid support to permit analysis of the oligonucleotides by solution-phase methods. *See, e.g., Wells et al. (1998) "Cleavage and Analysis of Material from Single Resin Beads," J. Org. Chem. 63:6430, which is incorporated by reference.*

20 [0039] A "character" when used in reference to a character of a character string refers to a subunit of the string. In one embodiment, the character of a character string encodes one subunit of an encoded biological molecule. Thus, for example, where the encoded biological molecule is a polynucleotide or oligonucleotide, a character of the string encodes a single nucleotide.

25 [0040] A "character string" is any entity capable of storing sequence information (e.g., the subunit structure of a biological molecule such as the nucleotide sequence of a nucleic acid, etc.). In one embodiment, the character string can be a simple sequence of characters (letters, numbers, or other symbols) or it can be a numeric or coded representation of such information in tangible or intangible (e.g., 30 electronic, magnetic, etc.) form. The character string need not be "linear," but can also exist in a number of other forms, e.g., a linked list or other non-linear array (e.g., used as a code to generate a linear array of characters), or the like. Character strings are

typically those which encode oligonucleotide or polynucleotide strings, directly or indirectly, including any encrypted strings, or images, or arrangements of objects which can be transformed unambiguously to character strings representing sequences of monomers or multimers in polynucleotides, or the like (whether made of natural or 5 artificial monomers).

[0041] The term “complement thereof” refers to nucleic acid that is both the same length as, and exactly complementary to, a given nucleic acid.

[0042] A “composition” refers to a combination of two or more different components. In certain embodiments, for example, a composition includes a solid 10 support that comprises one or more oligonucleotides, e.g., covalently or non-covalently attached to a surface of the support. In other embodiments, a composition includes one or more oligonucleotides in solution.

[0043] The term “deletion” in the context of a nucleic acid sequence refers to an alteration in which at least one nucleotide is removed from the nucleic acid 15 sequence, e.g., from a 5'-terminus, from a 3'-terminus, and/or from an internal position of the nucleic acid sequence.

[0044] The term “derivative” refers to a chemical substance related structurally to another substance, or a chemical substance that can be made from another substance (i.e., the substance it is derived from), e.g., through chemical or enzymatic 20 modification. To illustrate, oligonucleotides are optionally conjugated with biotin or a biotin derivative. To further illustrate, one nucleic acid can be “derived” from another through processes, such as chemical synthesis based on knowledge of the sequence of the other nucleic acid, amplification of the other nucleic acid, or the like.

[0045] The term “detectably bind” refers to binding between at least two 25 molecular species (e.g., a probe nucleic acid and a target nucleic acid, a sequence specific antibody and a target nucleic acid, etc.) that is detectable above a background signal (e.g., noise) using one or more methods of detection.

[0046] Nucleic acids are “extended” or “elongated” when additional 30 nucleotides (or other analogous molecules) are incorporated into the nucleic acids. For example, a nucleic acid is optionally extended by a nucleotide incorporating

biocatalyst, such as a polymerase that typically adds nucleotides at the 3' terminal end of a nucleic acid.

[0047] An "extended primer nucleic acid" refers to a primer nucleic acid to which one or more additional nucleotides have been added or otherwise incorporated (e.g., covalently bonded thereto).

[0048] Nucleic acids "hybridize" or "bind" when they associate with one another, typically in solution. Nucleic acids hybridize due to a variety of well characterized physico-chemical forces, such as hydrogen bonding, solvent exclusion, base stacking and the like. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes part I chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays," (Elsevier, New York), as well as in Ausubel (Ed.) Current Protocols in Molecular Biology, Volumes I, II, and III, 1997, which is incorporated by reference. Hames and Higgins (1995) Gene Probes 1 IRL Press at Oxford University Press, Oxford, England, (Hames and Higgins 1) and Hames and Higgins (1995) Gene Probes 2 IRL Press at Oxford University Press, Oxford, England (Hames and Higgins 2) provide details on the synthesis, labeling, detection and quantification of DNA and RNA, including oligonucleotides. Both Hames and Higgins 1 and 2 are incorporated by reference.

[0049] "Stringent hybridization wash conditions" in the context of nucleic acid hybridization assays or experiments, such as nucleic acid amplification reactions, Southern and northern hybridizations, or the like, are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993), *supra*. and in Hames and Higgins, 1 and 2.

[0050] For purposes of the present invention "highly stringent" hybridization and wash conditions are selected to be at least about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the test sequence hybridizes to a perfectly matched oligonucleotide. Very stringent conditions are selected to be equal to the  $T_m$  for a particular oligonucleotide.

[0051] An example of stringent hybridization conditions for hybridization of complementary nucleic acids on a filter in a Southern or northern blot is 50% formalin with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (see,

5 Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001), which is incorporated by reference, for a description of SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove background probe signal. An example low stringency wash is 2x SSC at 40°C for 15 minutes. In general, a signal to noise ratio of  
10 5x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

[0052] Comparative hybridization can be used to identify nucleic acids of the invention.

[0053] In particular, detection of stringent hybridization in the context of the  
15 present invention indicates strong structural similarity to, e.g., the nucleic acids provided in the sequence listing herein. For example, it is desirable to identify test nucleic acids that hybridize to the exemplar nucleic acids herein under stringent conditions. One measure of stringent hybridization is the ability to detectably hybridize to one of the listed nucleic acids (e.g., nucleic acids with sequences selected from SEQ  
20 ID NOS: 1-12 and 15-24 and complements thereof) under stringent conditions. Stringent hybridization and wash conditions can easily be determined empirically for any test nucleic acid.

[0054] For example, in determining highly stringent hybridization and wash conditions, the stringency of the hybridization and wash conditions are gradually  
25 increased (e.g., by increasing temperature, decreasing salt concentration, increasing detergent concentration and/or increasing the concentration of organic solvents such as formalin in the hybridization or wash), until a selected set of criteria is met. For example, the stringency of the hybridization and wash conditions are gradually increased until a oligonucleotide consisting of or comprising one or more nucleic acid  
30 sequences selected from SEQ ID NOS: 1-12 and 15-24 and complementary polynucleotide sequences thereof, binds to a perfectly matched complementary target (again, a nucleic acid comprising one or more nucleic acid sequences selected from

SEQ ID NOS: 1-12 and 15-24 and complementary polynucleotide sequences thereof), with a signal to noise ratio that is at least 5x as high as that observed for hybridization of the oligonucleotide to non-target nucleic acids present in the sample (e.g., nucleic acids from organisms other than SARS-CoV).

5 [0055] The detection of target nucleic acids which hybridize to the nucleic acids represented by SEQ ID NOS: 1-12 and 15-24 under high stringency conditions are a feature of the invention. Examples of such nucleic acids include those with one or a few silent or conservative nucleic acid substitutions as compared to a given nucleic acid sequence.

10 [0056] The terms "identical" or percent "identity" in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, e.g., as measured using one of the sequence comparison algorithms available to persons of skill or by visual inspection. Exemplary algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST programs, which are described in, e.g., Altschul et al. (1990) "Basic local alignment search tool" J. Mol. Biol. 215:403-410, Gish et al. (1993) "Identification of protein coding regions by database similarity search" Nature Genet. 3:266-272,

15 20 Madden et al. (1996) "Applications of network BLAST server" Meth. Enzymol. 266:131-141, Altschul et al. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs" Nucleic Acids Res. 25:3389-3402, and Zhang et al. (1997) "PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation" Genome Res. 7:649-656, which are each incorporated by reference. Many other optimal alignment algorithms are also known in the art and are optionally utilized to determine percent sequence identity.

25 [0057] The phrase "in solution" refers to an assay or reaction condition in which the components of the assay or reaction are not attached to a solid support and are present in a liquid medium. Exemplary liquid mediums include aqueous and organic fluids. For example, certain assays of the invention include incubating oligonucleotides together with SARS-CoV nucleic acids and SARS-CoV nucleic acid amplicons in solution to allow hybridization to occur.

[0058] The term "insertion" in the context of a nucleic acid sequence refers to an alteration in which at least one nucleotide is added to the nucleic acid sequence, e.g., at a 5'-terminus, at a 3'-terminus, and/or at an internal position of the nucleic acid sequence.

5 [0059] A "label" refers to a moiety attached (covalently or non-covalently), or capable of being attached, to a molecule, which moiety provides or is capable of providing information about the molecule (e.g., descriptive, identifying, etc. information about the molecule) or another molecule with which the labeled molecule interacts (e.g., hybridizes, etc.). Exemplary labels include fluorescent labels (including, e.g., quenchers or absorbers), weakly fluorescent labels, non-fluorescent labels, colorimetric labels, chemiluminescent labels, bioluminescent labels, radioactive labels, mass-modifying groups, antibodies, antigens, biotin, haptens, enzymes (including, e.g., peroxidase, phosphatase, etc.), and the like.

10 [0060] A "linker" refers to a chemical moiety that covalently or non-covalently attaches a compound or substituent group to another moiety, e.g., a probe nucleic acid, a primer nucleic acid, an amplicon, a solid support, or the like. For example, linkers are optionally used to attach oligonucleotides to a solid support (e.g., in a linear or other logic probe array). To further illustrate, a linker optionally attaches a label (e.g., a fluorescent dye, a radioisotope, etc.) to a probe nucleic acid, a primer nucleic acid, or the like. Linkers are typically at least bifunctional chemical moieties and in certain embodiments, they comprise cleavable attachments, which can be cleaved by, e.g., heat, an enzyme, a chemical agent, electromagnetic radiation, etc. to release materials or compounds from, e.g., a solid support. A careful choice of linker allows cleavage to be performed under appropriate conditions compatible with the stability of the compound and assay method. Generally a linker has no specific biological activity other than to, e.g., join chemical species together or to preserve some minimum distance or other spatial relationship between such species. However, the constituents of a linker may be selected to influence some property of the linked chemical species such as three-dimensional conformation, net charge, hydrophobicity, etc. Exemplary 15 linkers include, e.g., oligopeptides, oligonucleotides, oligopolyamides, oligoethyleneglycerols, oligoacrylamides, alkyl chains, or the like. Additional 20 description of linker molecules is provided in, e.g., Hermanson, Bioconjugate 25 30

Techniques, Elsevier Science (1996), Lytle et al. (1996) Nucleic Acids Res. 24(14):2793, Shchepino et al. (2001) Nucleosides, Nucleotides, & Nucleic Acids 20:369, Doronina et al (2001) Nucleosides, Nucleotides, & Nucleic Acids 20:1007, Trawick et al. (2001) Bioconjugate Chem. 12:900, Olejnik et al. (1998) Methods in Enzymology 291:135, and Pljevaljcic et al. (2003) J. Am. Chem. Soc. 125(12):3486, all of which are incorporated by reference.

5 [0061] A "mass modifying" group modifies the mass, typically measured in terms of molecular weight as daltons, of a molecule that comprises the group. For example, mass modifying groups that increase the discrimination between at least two 10 nucleic acids with single base differences in size or sequence can be used to facilitate sequencing using, e.g., molecular weight determinations.

10 [0062] A "mixture" refers to a combination of two or more different components. A "reaction mixture" refers a mixture that comprises molecules that can participate in and/or facilitate a given reaction. An "amplification reaction mixture" 15 refers to a solution containing reagents necessary to carry out an amplification reaction, and typically contains primers, a thermostable DNA polymerase, dNTP's, and a divalent metal cation in a suitable buffer. A reaction mixture is referred to as complete if it contains all reagents necessary to carry out the reaction, and incomplete if it 20 contains only a subset of the necessary reagents. It will be understood by one of skill in the art that reaction components are routinely stored as separate solutions, each containing a subset of the total components, for reasons of convenience, storage stability, or to allow for application-dependent adjustment of the component concentrations, and, that reaction components are combined prior to the reaction to 25 create a complete reaction mixture. Furthermore, it will be understood by one of skill in the art that reaction components are packaged separately for commercialization and that useful commercial kits may contain any subset of the reaction components, which includes the modified primers of the invention.

30 [0063] A "modified primer nucleic acid" refers to a primer nucleic acid that comprises a moiety or sequence of nucleotides that provides a desired property to the primer nucleic acid. In certain embodiments, for example, modified primer nucleic acids comprise "nucleic acid amplification specificity altering modifications" that, e.g., reduce non-specific nucleic acid amplification (e.g., minimize primer dimer

formation or the like), increase the yield of an intended target amplicon, and/or the like. Examples of nucleic acid amplification specificity altering modifications are described in, e.g., U.S. Pat. No. 6,001,611, entitled "MODIFIED NUCLEIC ACID AMPLIFICATION PRIMERS," issued December 14, 1999 to Will, which is

5 incorporated by reference. Other exemplary primer nucleic acid modifications include a "restriction site linker modification" in which a nucleotide sequence comprising a selected restriction site is attached, e.g., at 5'-terminus of a primer nucleic acid. Restriction site linkers are typically attached to primer nucleic acids to facilitate subsequent amplicon cloning or the like.

10 [0064] A "moiety" or "group" refers to one of the portions into which something, such as a molecule, is divided (e.g., a functional group, substituent group, or the like). For example, an oligonucleotide optionally comprises a quencher moiety, a labeling moiety, or the like.

[0065] The term "nucleic acid" refers to nucleotides (e.g., ribonucleotides, deoxyribonucleotides, dideoxynucleotides, etc.) and polymers that comprise such nucleotides covalently linked together, either in a linear or branched fashion. Exemplary nucleic acids include deoxyribonucleic acids (DNAs), ribonucleic acids (RNAs), DNA-RNA hybrids, oligonucleotides, polynucleotides, genes, cDNAs, aptamers, antisense nucleic acids, interfering RNAs (RNAis), molecular beacons, 20 nucleic acid probes, peptide nucleic acids (PNAs), locked nucleic acids (LNA<sup>TM</sup>s), PNA-DNA conjugates, PNA-RNA conjugates, LNA<sup>TM</sup>-DNA conjugates, LNA<sup>TM</sup>-RNA conjugates, etc.

[0066] A nucleic acid is typically single-stranded or double-stranded and will generally contain phosphodiester bonds, although in some cases, as outlined herein, 25 nucleic acid analogs are included that may have alternate backbones, including, for example and without limitation, phosphoramide (Beaucage et al. (1993) *Tetrahedron* 49(10):1925 and references therein; Letsinger (1970) *J. Org. Chem.* 35:3800; Sprinzl et al. (1977) *Eur. J. Biochem.* 81:579; Letsinger et al. (1986) *Nucl. Acids Res.* 14: 3487; Sawai et al. (1984) *Chem. Lett.* 805; Letsinger et al. (1988) *J. Am. Chem. Soc.* 30 110:4470; and Pauwels et al. (1986) *Chemica Scripta* 26: 1419, which are each incorporated by reference), phosphorothioate (Mag et al. (1991) *Nucleic Acids Res.* 19:1437; and U.S. Pat. No. 5,644,048, which are both incorporated by reference),

phosphorodithioate (Bru et al. (1989) J. Am. Chem. Soc. 111:2321, which is incorporated by reference), O-methylphosphoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press (1992), which is incorporated by reference), and peptide nucleic acid backbones and linkages (see, Egholm (1992) J. Am. Chem. Soc. 114:1895; Meier et al. (1992) Chem. Int. Ed. Engl. 31:1008; Nielsen (1993) Nature 365:566; and Carlsson et al. (1996) Nature 380:207, which are each incorporated by reference). Other analog nucleic acids include those with positively charged backbones (Denpcy et al. (1995) Proc. Natl. Acad. Sci. USA 92:6097, which is incorporated by reference); non-ionic backbones (U.S. Pat. Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Angew (1991) Chem. Intl. Ed. English 30: 423; Letsinger et al. (1988) J. Am. Chem. Soc. 110:4470; Letsinger et al. (1994) Nucleoside & Nucleotide 13:1597; Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y. S. Sanghvi and P. Dan Cook; Mesmaeker et al. (1994) Bioorganic & Medicinal Chem. Lett. 4: 395; Jeffs et al. (1994) J. Biomolecular NMR 34:17; and Tetrahedron Lett. 37:743 (1996), which are each incorporated by reference) and non-ribose backbones, including those described in U.S. Pat. Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, Carbohydrate Modifications in Antisense Research, Ed. Y. S. Sanghvi and P. Dan Cook, which references are each incorporated by reference. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins et al. (1995) Chem. Soc. Rev. pp169-176, which is incorporated by reference). Several nucleic acid analogs are also described in, e.g., Rawls, C & E News Jun. 2, 1997 page 35, which is incorporated by reference. These modifications of the ribose-phosphate backbone may be done to facilitate the addition of additional moieties such as labels, or to alter the stability and half-life of such molecules in physiological environments.

[0067] In addition to these naturally occurring heterocyclic bases that are typically found in nucleic acids (e.g., adenine, guanine, thymine, cytosine, and uracil), nucleic acid analogs also include those having non-naturally occurring heterocyclic or modified bases, many of which are described, or otherwise referred to, herein. In particular, many non-naturally occurring bases are described further in, e.g., Seela et al. (1991) Helv. Chim. Acta 74:1790, Grein et al. (1994) Bioorg. Med. Chem. Lett. 4:971-

976, and Seela et al. (1999) *Helv. Chim. Acta* 82:1640, which are each incorporated by reference. To further illustrate, certain bases used in nucleotides that act as melting temperature ( $T_m$ ) modifiers are optionally included. For example, some of these include 7-deazapurines (e.g., 7-deazaguanine, 7-deazaadenine, etc.), pyrazolo[3,4-d]pyrimidines, propynyl-dN (e.g., propynyl-dU, propynyl-dC, etc.), and the like. See, e.g., U.S. Pat. No. 5,990,303, entitled "SYNTHESIS OF 7-DEAZA-2'-DEOXYGUANOSINE NUCLEOTIDES," which issued November 23, 1999 to Seela, which is incorporated by reference. Other representative heterocyclic bases include, e.g., hypoxanthine, inosine, xanthine; 8-aza derivatives of 2-aminopurine, 2,6-diaminopurine, 2-amino-6-chloropurine, hypoxanthine, inosine and xanthine; 7-deaza-8-aza derivatives of adenine, guanine, 2-aminopurine, 2,6-diaminopurine, 2-amino-6-chloropurine, hypoxanthine, inosine and xanthine; 6-azacytosine; 5-fluorocytosine; 5-chlorocytosine; 5-iodocytosine; 5-bromocytosine; 5-methylcytosine; 5-propynylcytosine; 5-bromovinyluracil; 5-fluorouracil; 5-chlorouracil; 5-iodouracil; 5-bromouracil; 5-trifluoromethyluracil; 5-methoxymethyluracil; 5-ethynyluracil; 5-propynyluracil, and the like.

[0068] Examples of modified bases and nucleotides are also described in, e.g., U.S. Pat. No. 5,484,908, entitled "OLIGONUCLEOTIDES CONTAINING 5-PROPYNYL PYRIMIDINES," issued January 16, 1996 to Froehler et al., U.S. Pat. No. 5,645,985, entitled "ENHANCED TRIPLE-HELIX AND DOUBLE-HELIX FORMATION WITH OLIGOMERS CONTAINING MODIFIED PYRIMIDINES," issued July 8, 1997 to Froehler et al., U.S. Pat. No. 5,830,653, entitled "METHODS OF USING OLIGOMERS CONTAINING MODIFIED PYRIMIDINES," issued November 3, 1998 to Froehler et al., U.S. Pat. No. 6,639,059, entitled "SYNTHESIS OF [2.2.1]BICYCLO NUCLEOSIDES," issued October 28, 2003 to Kochkine et al., U.S. Pat. No. 6,303,315, entitled "ONE STEP SAMPLE PREPARATION AND DETECTION OF NUCLEIC ACIDS IN COMPLEX BIOLOGICAL SAMPLES," issued October 16, 2001 to Skoulios, and U.S. Pat. Application Pub. No. 2003/0092905, entitled "SYNTHESIS OF [2.2.1]BICYCLO NUCLEOSIDES," by Kochkine et al. that published May 15, 2003, which are each incorporated by reference.

[0069] The term "nucleic acid detection reagent" refers to a reagent that detectably binds (e.g., hydrogen bonds in nucleic acid hybridization, in antibody-

antigen recognition, or the like, or other types of binding interactions) to a SARS-CoV nucleic acid. For example, nucleic acids (e.g., probe nucleic acids, primer nucleic acids, etc.) that comprise sequences selected from SEQ ID NOS: 1-12 and 15-24 or complements thereof bind to SARS-CoV nucleic acids. Other exemplary nucleic acid 5 detection reagents include sequence specific antibodies that specifically bind to SARS-CoV nucleic acids.

[0070] A "nucleotide" refers to an ester of a nucleoside, e.g., a phosphate ester of a nucleoside. For example, a nucleotide can include 1, 2, 3, or more phosphate groups covalently linked to a 5' position of a sugar moiety of the nucleoside.

10 [0071] A "nucleotide incorporating biocatalyst" refers to a catalyst that catalyzes the incorporation of nucleotides into a nucleic acid. Nucleotide incorporating biocatalysts are typically enzymes. An "enzyme" is a protein- and/or nucleic acid-based catalyst that acts to reduce the activation energy of a chemical reaction involving other compounds or "substrates." A "nucleotide incorporating enzyme" refers to an 15 enzyme that catalyzes the incorporation of nucleotides into a nucleic acid, e.g., during nucleic acid amplification or the like. Exemplary nucleotide incorporating enzymes include, e.g., polymerases, terminal transferases, reverse transcriptases, telomerases, polynucleotide phosphorylases, and the like.

20 [0072] An "oligonucleotide" refers to a nucleic acid that includes at least two nucleic acid monomer units (e.g., nucleotides), typically more than three monomer units, and more typically greater than ten monomer units. The exact size of an oligonucleotide generally depends on various factors, including the ultimate function or use of the oligonucleotide. Oligonucleotides are optionally prepared by any suitable method; including, but not limited to, isolation of an existing or natural sequence, DNA 25 replication or amplification, reverse transcription, cloning and restriction digestion of appropriate sequences, or direct chemical synthesis by a method such as the phosphotriester method of Narang et al. (1979) Meth. Enzymol. 68:90-99; the phosphodiester method of Brown et al. (1979) Meth. Enzymol. 68:109-151; the diethylphosphoramide method of Beaucage et al. (1981) Tetrahedron Lett. 22:1859-30 1862; the triester method of Matteucci et al. (1981) J. Am. Chem. Soc. 103:3185-3191; automated synthesis methods; or the solid support method of U.S. Pat. No. 4,458,066,

or other methods known in the art. All of these references are incorporated by reference.

[0073] The term “probe nucleic acid” or “probe” refers to a labeled or unlabeled oligonucleotide capable of selectively hybridizing to a target nucleic acid under suitable conditions. Typically, a probe is sufficiently complementary to a specific target sequence contained in a SARS-CoV nucleic acid sample to form a stable hybridization duplex with the target sequence under a selected hybridization condition, such as, but not limited to, a stringent hybridization condition. A hybridization assay carried out using the probe under sufficiently stringent hybridization conditions permits the selective detection of a specific target sequence. The term “hybridizing region” refers to that region of a nucleic acid that is exactly or substantially complementary to, and therefore hybridizes to, the target sequence. For use in a hybridization assay for the discrimination of single nucleotide differences in sequence, the hybridizing region is typically from about 8 to about 100 nucleotides in length. Although the hybridizing region generally refers to the entire oligonucleotide, the probe may include additional nucleotide sequences that function, for example, as linker binding sites to provide a site for attaching the probe sequence to a solid support or the like. In certain embodiments, a probe of the invention comprises one or more labels (e.g., a reporter dye, a quencher moiety, etc.), such as a 5'-nuclease probe, a FRET probe, a molecular beacon, or the like, which can also be utilized to detect hybridization between the probe and target nucleic acids in a sample. In some embodiments, the hybridizing region of the probe is completely complementary to the target sequence. However, in general, complete complementarity is not necessary; stable duplexes may contain mismatched bases or unmatched bases. Modification of the stringent conditions may be necessary to permit a stable hybridization duplex with one or more base pair mismatches or unmatched bases. Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001), which is incorporated by reference, provides guidance for suitable modification. Stability of the target/probe duplex depends on a number of variables including length of the oligonucleotide, base composition and sequence of the oligonucleotide, temperature, and ionic conditions. One of skill in the art will recognize that, in general, the exact complement of a given probe is similarly useful as a probe. Exemplary probes of the

invention that bind to SARS-CoV nucleic acids include oligonucleotides that comprise sequences selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. One of skill in the art will also recognize that, in certain embodiments, probe nucleic acids can also be used as primer nucleic acids.

[0074] A "primer nucleic acid" or "primer" is a nucleic acid that can hybridize to a template nucleic acid (e.g., a SARS-CoV nucleic acid) and permit chain extension or elongation using, e.g., a nucleotide incorporating biocatalyst, such as a polymerase under appropriate reaction conditions. A primer nucleic acid is typically a natural or synthetic oligonucleotide (e.g., a single-stranded oligodeoxyribonucleotide, etc.). Although other primer nucleic acid lengths are optionally utilized, they typically comprise hybridizing regions that range from about 8 to about 100 nucleotides in length. Short primer nucleic acids generally utilize cooler temperatures to form sufficiently stable hybrid complexes with template SARS-CoV nucleic acid. A primer nucleic acid that is at least partially complementary to a subsequence of a template SARS-CoV nucleic acid is typically sufficient to hybridize with the template for extension to occur. A primer nucleic acid can be labeled, if desired, by incorporating a label detectable by, e.g., spectroscopic, photochemical, biochemical, immunochemical, chemical, or other techniques. To illustrate, useful labels include radioisotopes, fluorescent dyes, electron-dense reagents, enzymes (as commonly used in ELISAs), biotin, or haptens and proteins for which antisera or monoclonal antibodies are available. Many of these and other labels are described further herein and/or are otherwise known in the art. Exemplary primer nucleic acids of the invention that bind to SARS-CoV nucleic acids include oligonucleotides that comprise sequences selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. One of skill in the art will recognize that, in certain embodiments, primer nucleic acids can also be used as probe nucleic acids.

[0075] A "quencher moiety" or "quencher" refers to a moiety that reduces and/or is capable of reducing the detectable emission of radiation, e.g., fluorescent or

luminescent radiation, from a source that would otherwise have emitted this radiation. A quencher typically reduces the detectable radiation emitted by the source by at least 50%, typically by at least 80%, and more typically by at least 90%. Exemplary quenchers are provided in, e.g., U.S. Pat. No. 6,465,175, entitled

5 "OLIGONUCLEOTIDE PROBES BEARING QUENCHABLE FLUORESCENT LABELS, AND METHODS OF USE THEREOF," which issued October 15, 2002 to Horn et al., which is incorporated by reference.

[0076] The term "sample" refers to any substance containing or presumed to contain one or more SARS-CoV nucleic acids including, but not limited to, clinical 10 samples (e.g., tissue or fluid isolated from one or more subjects or individuals), in vitro cell culture constituents, environmental samples, and the like. Exemplary samples include blood, plasma, serum, feces, bronchoalveolar lavage, nasal pharyngeal swabs and tissues, urine, synovial fluid, seminal fluid, seminal plasma, prostatic fluid, vaginal fluid, cervical fluid, uterine fluid, cervical scrapings, amniotic fluid, anal scrapings, 15 mucus, sputum, tissue, and the like.

[0077] The phrase "sample derived from a subject" refers to a sample obtained from the subject, whether or not that sample undergoes one or more processing steps (e.g., cell lysis, debris removal, stabilization, etc.) prior to analysis. To illustrate, samples can be derived from subjects by scraping, venipuncture, 20 swabbing, biopsy, or other techniques known in the art.

[0078] The term "severe acute respiratory syndrome coronavirus" or "SARS-CoV" refers to a positive-strand RNA-enveloped virus of the family *Coronaviridae* and the order *Nidovirales*, which causes the severe acute respiratory syndrome (SARS) in subjects infected with the virus. See, e.g., Ksiazek et al. (2003) "A 25 novel coronavirus associated with severe acute respiratory syndrome," *N. Engl. J. Med.* 348:1953-1966, Drosten et al. (2003) "Identification of a novel coronavirus in patients with severe acute respiratory syndrome," *N. Engl. J. Med.* 348:1967-1976, Marra et al. (2003) "The genome sequence of the SARS-associated coronavirus," *Science* 300:1399-1404, Ruan et al. (2003) "Comparative full-length genome sequence analysis 30 of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection," *Lancet* 361:1779-1785, and Cavanagh (2003) "Nidovirales: a new order comprising *Coronaviridae* and *Arteriviridae*," *Arch. Virol.* 14:629-633, which

are each incorporated by reference. Exemplary GenBank® accession numbers for complete sequences of the SARS-CoV genome are NC\_004718, AY283794, AY278489, AY278741, AY559097, AY559096, AY559095, AY559094, AY559093, AY559092, AY559091, AY559090, AY559089, AY559088, AY559087, AY559086, 5 AY559085, AY559084, AY559083, AY559082, AY559081, AY274119, AY323977, AY291315, and the like. To further illustrate, Figure 1 provides the nucleotide sequence of the SARS-CoV genome (Urbani strain) as determined by the Centers for Disease Control.

10 [0079] The term “severe acute respiratory syndrome coronavirus nucleic acid” or “SARS-CoV nucleic acid” refers to a nucleic acid that is derived or isolated from a SARS-CoV and/or an amplicon thereof.

15 [0080] The term “selectively bind” or “selective binding” in the context of nucleic acid detection reagents refers to a nucleic acid detection reagent that binds to SARS-CoV nucleic acids to a greater extent than the nucleic acid detection reagent binds, under the same hybridization conditions, to non-target nucleic acids.

[0081] The term “selectively detect” refers to the ability to detect a SARS-CoV nucleic acid to a greater extent than nucleic acids from other organisms.

20 [0082] “Selectively hybridizing” or “selective hybridization” occurs when a nucleic acid sequence hybridizes to a specified nucleic acid target sequence to a detectably greater degree than its hybridization to non-target nucleic acid sequences. Selectively hybridizing sequences have at least 50%, or 60%, or 70%, or 80%, or 90% sequence identity or more, e.g., typically 95-100% sequence identity (i.e., complementarity) with each other.

25 [0083] A “sequence” of a nucleic acid refers to the order and identity of nucleotides in the nucleic acid. A sequence is typically read in the 5' to 3' direction.

[0084] A “sequence specific antibody” refers to an antibody that detectably binds to SARS-CoV nucleic acids.

30 [0085] A “sequencing reaction” refers to a reaction that includes, e.g., the use of terminator nucleotides and which is designed to elucidate the sequence of nucleotides in a given nucleic acid.

[0086] A "set" refers to a collection of at least two things. For example, a set may include 2, 3, 4, 5, 10, 20, 50, 100, 1,000 or other number of molecule or sequence types. For example, certain aspects of the invention include reaction mixtures having sets of amplicons. A "subset" refers to any portion of a set.

5 [0087] A "solid support" refers to a solid material that can be derivatized with, or otherwise attached to, a chemical moiety, such as an oligonucleotide or the like. Exemplary solid supports include plates, beads, microbeads, tubes, fibers, whiskers, combs, hybridization chips (including microarray substrates, such as those used in GeneChip® probe arrays (Affymetrix, Inc., Santa Clara, CA, USA) and the like), membranes, single crystals, ceramic layers, self-assembling monolayers, and the like.

10 [0088] An oligonucleotide is "specific" for a target sequence if the number of mismatches present between the oligonucleotide and the target sequence is less than the number of mismatches present between the oligonucleotide and non-target sequences that might be present in a sample. Hybridization conditions can be chosen under which 15 stable duplexes are formed only if the number of mismatches present is no more than the number of mismatches present between the oligonucleotide and the target sequence. Under such conditions, the target-specific oligonucleotide can form a stable duplex only with a target sequence. Thus, the use of target-specific primers under suitably stringent amplification conditions enables the specific amplification of those sequences, 20 which contain the target primer binding sites. Similarly, the use of target-specific probes under suitably stringent hybridization conditions enables the detection of a specific target sequence.

15 [0089] A test nucleic acid is said to "specifically hybridize" to an oligonucleotide when it hybridizes at least one-half as well to the oligonucleotide as to the perfectly matched complementary target, i.e., with a signal to noise ratio at least 25 one-half as high as hybridization of the oligonucleotide to the target under conditions in which the perfectly matched oligonucleotide binds to the perfectly matched complementary target with a signal to noise ratio that is at least about 5x-10x as high as that observed for hybridization to non-target nucleic acids present in a sample.

[0090] A "subject" refers to an organism. Typically, the organism is a mammalian organism, particularly a human organism. In certain embodiments, for example, a subject is a patient suspected of having a SARS-CoV infection.

5 [0091] A "subsequence" or "segment" refers to any portion of an entire nucleic acid sequence.

[0092] A "substantially identical variant" in the context of nucleic acids or polypeptides, refers to two or more sequences that have at least 85%, typically at least 90%, more typically at least 95% nucleotide or sequence identity to one another when compared and aligned for maximum correspondence, as measured using, e.g., a sequence comparison algorithm or by visual inspection. The substantial identity generally exists over a region of the sequences that is at least about 15 nucleotides or amino acids in length, more typically over a region that is at least about 20 nucleotides or amino acids in length, and even more typically the sequences are substantially identical over a region of at least about 25 nucleotides or amino acids in length. In 10 some embodiments, for example, the sequences are substantially identical over the entire length of the nucleic acids or polypeptides being compared.

15 [0093] The term "substitution" in the context of a nucleic acid sequence refers to an alteration in which at least one nucleotide of the nucleic acid sequence is replaced by a different nucleotide.

20 [0094] The terms "target sequence," "target region," and "target nucleic acid" refer to a region of a nucleic acid, which is to be amplified, detected, or otherwise analyzed.

25 [0095] A "terminator nucleotide" refers to a nucleotide, which upon incorporation into a nucleic acid prevents further extension of the nucleic acid, e.g., by at least one nucleotide incorporating biocatalyst.

30 [0096] A "thermostable enzyme" refers to an enzyme that is stable to heat, is heat resistant and retains sufficient catalytic activity when subjected to elevated temperatures for selected periods of time. For example, a thermostable polymerase retains sufficient activity to effect subsequent primer extension reactions when subjected to elevated temperatures for the time necessary to effect denaturation of double-stranded nucleic acids. Heating conditions necessary for nucleic acid

denaturation are well known in the art and are exemplified in U.S. Pat. Nos. 4,683,202 and 4,683,195, which are both incorporated by reference. As used herein, a thermostable polymerase is typically suitable for use in a temperature cycling reaction such as the polymerase chain reaction ("PCR"). For a thermostable polymerase, 5 enzymatic activity refers to the catalysis of the combination of the nucleotides in the proper manner to form primer extension products that are complementary to a template nucleic acid (e.g., SARS-CoV nucleic acids).

## II. OVERVIEW

[0097] The Centers for Disease Control and Prevention (CDC) has sequenced 10 the genome for a coronavirus believed to be responsible for the global epidemic of severe acute respiratory syndrome or SARS. Figure 1 shows this nucleotide sequence of the SARS-CoV genome, which corresponds to GenBank® accession number AY278741. *See also*, Rota et al. (2003) "Characterization of a novel coronavirus associated with severe acute respiratory syndrome," Science 300(5624):1394-1399, 15 which is incorporated by reference. The sequence data confirm that the SARS coronavirus is a previously unrecognized coronavirus. All of the sequence, except for the leader sequence, was derived directly from viral RNA. The genome of the SARS coronavirus sequenced by the CDC is 29,727 nucleotides in length and the genome organization is similar to that of other coronaviruses. *See, e.g.*, Drosten et al. (2003) 20 "Identification of a novel coronavirus in patients with severe acute respiratory syndrome," N. Engl. J. Med. 348:1967-1976, which is incorporated by reference. Open reading frames corresponding to the polymerase protein (polymerase 1a, 1b), spike protein (S), small membrane protein (E), membrane protein (M) and nucleocapsid protein (N) have been identified.

[0098] The invention relates to the selective detection of SARS-CoV. In 25 particular, based on new detection strategies utilizing selected target regions of the SARS-CoV genome, SARS-CoV infections can be diagnosed using the methods and reagents described herein. In certain embodiments, for example, certain nucleic acid detection reagents described herein target a region of the SARS-CoV genome that 30 encodes an RNA-dependent RNA polymerase (RdRp), which has been shown to be stable against mutations compared to other regions of the SARS-CoV genome. The

nucleic acid detection reagents described herein generally detectably bind, under selected assay conditions, to nucleotide sequences that are present in SARS-CoV, but which are not present in other species, thereby minimizing the occurrence of, e.g., false positives. The detection of SARS-CoV nucleic acids with certain nucleic acid 5 detection reagents of the invention is illustrated in, for example, in the examples provided below. Many other features of the invention are also described herein.

[0099] To further illustrate, certain methods of the invention include contacting or incubating nucleic acid detection reagents with nucleic acids in or from samples derived from subjects (e.g., human patients suspected of having SARS-CoV infections, 10 etc.). In certain embodiments, target regions of the nucleic acids in the samples are amplified prior to or simultaneously with being contacted with the nucleic acid detection reagents. These methods also include monitoring (e.g., at a single time point, at multiple discrete time points, continuously over a selected time period, etc.) binding between the nucleic acids and/or amplicons, and the nucleic acid detection reagents to 15 determine whether the SARS-CoV is present in the samples, e.g., to diagnose patients from which the samples were derived, to monitor courses of treatment for patients diagnosed with SARS infections, and/or the like.

[0100] Other methods of the invention include contacting or incubating nucleic acids from samples with at least one pair of primer nucleic acids that include at least 20 one nucleic acid having a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof in which the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, in nucleic acid amplification reactions. In addition, these methods also include detecting amplicons 25 during or after the amplification reactions are performed to detect whether the SARS-CoV is present in the samples. These methods are also optionally repeated at selected time points.

[0101] In addition to compositions and reaction mixtures, the invention also relates to kits and systems for detecting SARS-CoV, and to related computers and 30 computer readable media.

### III. NUCLEIC ACID DETECTION REAGENTS

[0102] The nucleic acid detection reagents of the invention include various

embodiments, including probe nucleic acids, primer nucleic acids, and sequence 5 specific antibodies. In some embodiments, for example, certain nucleic acid detection reagents described herein target a segment of the SARS-CoV genome that encodes an RdRp. This region of the SARS-CoV genome is also referred to in, e.g., Ng et. (2004) "Detection of Severe Acute Respiratory Syndrome Coronavirus in Blood of Infected Patients" J. Clin. Microbiol. 42:347-350, which is incorporated by reference.

[0103] Exemplary nucleic acid detection reagents that target SARS-CoV

10 nucleic acids include oligonucleotides comprising sequences selected from SEQ ID NOS: 1-12 and 15-24 or complements thereof. SEQ ID NOS: 1-12 and 15-24 are shown in Table I.

TABLE I

SEQ ID NO: 1	5'- TGATGGTGTGTTCCACTCAACATC -3'
SEQ ID NO: 2	5'- ACAGTGGTACAGGACAGGGCAATTAC -3'
SEQ ID NO: 3	5'- GGTGTAAGTGCAGCCCGTCTACAC -3'
SEQ ID NO: 4	5'- CGAATTTCGCTCACAGCATAAATG -3'
SEQ ID NO: 5	5'- ATTGGAGAGTACACCTTGAAAAAG -3'
SEQ ID NO: 6	5'- GCTAGACTTCGTGCAAAACACTACG -3'
SEQ ID NO: 7	5'- TGTGGCTAGTTGTGATGCTATCATG -3'
SEQ ID NO: 8	5'- GCTACACATCACGATAAATTCACTG -3'
SEQ ID NO: 9	5'- CACTCAAATCTGCTACGTGTATTAC -3'
SEQ ID NO: 10	5'- CTCGCTATGGATGAATTACATACAGC -3'
SEQ ID NO: 11	5'- GGTTGGGATTATCCAAAATGTGA -3'
SEQ ID NO: 12	5'- GGCATCATCAGAAAGAACATCAT -3'
SEQ ID NO: 15	5'- ACTATATGTTAAACCAGGGTGG -3'
SEQ ID NO: 16	5'- ATTTACATTGGCTGTAAACAGC -3'
SEQ ID NO: 17	5'- AGCTAACGAGTGTGCGCAAGTATAAGTGAGATG -3'
SEQ ID NO: 18	5'- CCTCTCTTGTCTGTGCTCGCAAAC -3'
SEQ ID NO: 19	5'- AGAACAAAGAGAGGCCATTATCCTAAG -3'
SEQ ID NO: 20	5'- TTAACATATAGTGAGCCGCCACAC -3'
SEQ ID NO: 21	5'- AGAGCCATGCTAACAT -3'
SEQ ID NO: 22	5'- GGTTGGGATTATCCAAAATGTGAC -3'
SEQ ID NO: 23	5'- GGTTGGGATTATCCAAAATGTGA -3'
SEQ ID NO: 24	5'- GTGTGGCGGCTCACTATATGTTA -3'

15 [0104] To further illustrate, Figure 2 shows alignments of nucleotide sequences from various SARS-CoV isolates and from certain non-target organisms relative to one another. As shown, Group 1 shows the alignment of non-target nucleotide sequences

from the Human Coronavirus (HCoV) 229E (Accession No. NC\_002645) and the Porcine epidemic diarrhea virus (PEDV) (Accession No. NC\_003436) with target nucleotide sequences from three different SARS-CoV isolates, namely, Toronto (Accession No. NC\_004718), Singapore (Accession No. AY283794), and Guangzhou (Accession No. AY278489). As shown, the positions of oligonucleotides with sequences corresponding to SEQ ID NOS: 23 and 24 are underlined. Group 2 shows the alignment of non-target nucleotide sequences from the Bovine Coronavirus (BCoV) (Accession No. NC\_003045) and the Murine hepatitis virus (MHV) (Accession No. NC\_001846) with the target nucleotide sequences from the different SARS-CoV isolates shown in Group 1. The positions of oligonucleotides with sequences corresponding to SEQ ID NOS: 23 and 24 are also underlined in the alignment of Group 2. Group 3 illustrates the alignment of a non-target nucleotide sequence from the Avian Infectious Bronchitis Virus (IBV) (Accession No. NC\_001451) with the target nucleotide sequences from the different SARS-CoV isolates shown in Group 1. The positions of oligonucleotides with sequences corresponding to SEQ ID NOS: 23 and 24 are also underlined in the alignment of Group 3.

[0105] As mentioned above, nucleic acid detection reagents comprise oligonucleotides (e.g., probe nucleic acids, primer nucleic acids, etc.) in certain embodiments of the invention. Although other lengths are optionally utilized, oligonucleotides generally comprise sequences that are typically between about 8 and about 100 nucleotides in length, more typically between about 10 and about 75 nucleotides in length, still more typically between about 12 and about 50 nucleotides in length, and even more typically between about 15 and about 35 nucleotides in length (e.g., about 20, about 25, or about 30 nucleotides in length). Methods of preparing oligonucleotides, such as nucleic acid synthesis, are described further below.

[0106] Various approaches can be utilized by one of skill in the art to design oligonucleotides (e.g., substantially identical variants of nucleic acids having sequences selected from SEQ ID NOS: 1-12 and 15-24 or complements thereof) that selectively bind to SARS-CoV nucleic acids. To illustrate, the DNASTAR software package available from DNASTAR, Inc. (Madison, WI, USA) can be used for sequence alignments. For example, nucleic acid sequences for the SARS-CoV genome and non-

target sequences can be uploaded into DNASTar EditSeq program as individual files. To further illustrate, pairs of sequence files can be opened in the DNASTar MegAlign sequence alignment program and the Clustal W method of alignment can be applied. The parameters used for Clustal W alignments are optionally the default settings in the 5 software. MegAlign typically does not provide a summary of the percent identity between two sequences. This is generally calculated manually. From the alignments, regions having, e.g., less than 85% identity with one another are typically identified and oligonucleotide sequences in these regions can be selected. Many other sequence alignment algorithms and software packages are also optionally utilized. Sequence 10 alignment algorithms are also described in, e.g., Mount, Bioinformatics: Sequence and Genome Analysis, Cold Spring Harbor Laboratory Press (2001), and Durbin et al., Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press (1998), which are both incorporated by reference.

[0107] To further illustrate, optimal alignment of sequences for comparison can 15 be conducted, e.g., by the local homology algorithm of Smith & Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman & Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson & Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444, which are each incorporated by reference, 20 and by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (Madison, WI), or by even by visual inspection.

[0108] Another example algorithm that is suitable for determining percent 25 sequence identity is the BLAST algorithm, which is described in, e.g., Altschul et al. (1990) J. Mol. Biol. 215:403-410, which is incorporated by reference. Software for performing versions of BLAST analyses is publicly available through the National Center for Biotechnology Information on the world wide web at [ncbi.nlm.nih.gov/](http://ncbi.nlm.nih.gov/) as of April 20, 2004. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either 30 match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., *supra*). These initial neighborhood word hits act as

seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; 5 always  $> 0$ ) and N (penalty score for mismatching residues; always  $< 0$ ). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; 10 or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an 15 expectation (E) of 10, and the BLOSUM62 scoring matrix (see, Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915, which is incorporated by reference).

[0109] In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul (1993) Proc. Nat'l. Acad. Sci. USA 90:5873-5787, which 20 is incorporated by reference). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence (and, therefore, homologous) if the smallest sum probability in a comparison 25 of the test nucleic acid to the reference nucleic acid is less than about 0.1, or less than about 0.01, and or even less than about 0.001.

[0110] An additional example of a useful sequence alignment algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the 30 clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle (1987) J. Mol. Evol. 35:351-360,

which is incorporated by reference. The method used is similar to the method described by Higgins & Sharp (1989) CABIOS 5:151-153, which is incorporated by reference. The program can align, e.g., up to 300 sequences of a maximum length of 5,000 letters. The multiple alignment procedure begins with the pairwise alignment of 5 the two most similar sequences, producing a cluster of two aligned sequences. This cluster can then be aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences can be aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program can also be used to plot a 10 dendrogram or tree representation of clustering relationships. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison.

[0111] The probes and primers of the invention optionally include one or more labels, which are described further below. In addition, probes and primers optionally 15 include various other modifications, such as modified nucleotides that alter hybridization melting temperatures, restriction site linkers to facilitate amplicon cloning, modifier groups that increase the specificity of nucleic acid amplification reactions, and/or the like. For example, certain modified nucleotides that increase nucleic acid hybridization melting temperatures are optionally included to permit the 20 use of smaller probes and primers, such as those including between about 8 and about 14 nucleotides. Examples of these modified oligonucleotides include those having one or more LNA™ monomers. Nucleotide analogs such as these are described further in, e.g., U.S. Pat. No. 6,639,059, entitled “SYNTHESIS OF [2.2.1]BICYCLO 25 NUCLEOSIDES,” issued October 28, 2003 to Kochkine et al., U.S. Pat. No. 6,303,315, entitled “ONE STEP SAMPLE PREPARATION AND DETECTION OF NUCLEIC ACIDS IN COMPLEX BIOLOGICAL SAMPLES,” issued October 16, 2001 to Skouv, and U.S. Pat. Application Pub. No. 2003/0092905, entitled “SYNTHESIS OF [2.2.1]BICYCLO NUCLEOSIDES,” by Kochkine et al. that published May 15, 2003, which are each incorporated by reference. Oligonucleotides comprising LNA™ 30 monomers are commercially available through, e.g., Exiqon A/S (Vedbæk, DK). Additional probe and primer modifications are referred to herein, including in the definitions provided above.

[0112] In certain embodiments, the nucleic acid detection reagents utilized as described herein are sequence specific antibodies that target SARS-CoV nucleic acids. Antibodies suitable for use in these embodiments of invention may be prepared by conventional methodology and/or by genetic engineering. Antibody fragments may be 5 genetically engineered, e.g., from the variable regions of the light and/or heavy chains ( $V_H$  and  $V_L$ ), including the hypervariable regions, or from both the  $V_H$  and  $V_L$  regions. For example, the term "antibodies" as used herein includes polyclonal and monoclonal antibodies and biologically active fragments thereof including among other possibilities "univalent" antibodies (Glennie et al. (1982) *Nature* 295:712); Fab proteins including 10 Fab' and  $F(ab')_2$  fragments whether covalently or non-covalently aggregated; light or heavy chains alone, typically variable heavy and light chain regions ( $V_H$  and  $V_L$  regions), and more typically including the hypervariable regions (otherwise known as the complementarity determining regions (CDRs) of the  $V_H$  and  $V_L$  regions);  $F_c$  proteins; "hybrid" antibodies capable of binding more than one antigen; constant- 15 variable region chimeras; "composite" immunoglobulins with heavy and light chains of different origins; "altered" antibodies with improved specificity and other characteristics as prepared by standard recombinant techniques, by mutagenic techniques, or other directed evolutionary techniques known in the art.

[0113] The sequence specific antibodies utilized as described herein may be 20 labeled or unlabeled. Suitable labels include, e.g., radionuclides, enzymes, coenzymes, fluorescent dyes, chemiluminescent dyes, chromogens, enzyme substrates or co-factors, enzyme inhibitors, free radicals, and the like. Such labeled reagents may be used in a variety of well known assays, such as radioimmunoassays, enzyme immunoassays, e.g., ELISA, fluorescent immunoassays, and the like. See, e.g., U.S. Pat. Nos. 3,766,162; 25 3,791,932; 3,817,837; and 4,233,402, which are each incorporated by reference. Additional labels are described further herein.

[0114] In some embodiments, polypeptides, such the RdRp referred to above are targeted for detection. Many techniques for detecting proteins are known in the art. For example, various electrophoretic assays (e.g., SDS-PAGE or the like), 30 immunoassays, mass spectrometric assays (e.g., matrix assisted laser desorption/ionization (MALDI)-based analyses, surface enhanced laser

desorption/ionization (SELDI)-based assays, etc.), and/or other approaches can be used to detect proteins encoded by SARS-CoV nucleic acids. Many of these and other suitable protein detection methods are described in the references cited herein.

[0115] In practicing the present invention, many conventional techniques in molecular biology and recombinant DNA are optionally used. These techniques are well known and are explained in, for example, Current Protocols in Molecular Biology, Volumes I, II, and III, 1997 (F. M. Ausubel ed.); Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001; Berger and Kimmel, Guide to Molecular Cloning Techniques, 10 Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger), DNA Cloning: A Practical Approach, Volumes I and II, 1985 (D. N. Glover ed.); Oligonucleotide Synthesis, 1984 (M. L. Gait ed.); Nucleic Acid Hybridization, 1985, (Hames and Higgins); Transcription and Translation, 1984 (Hames and Higgins eds.); Animal Cell Culture, 1986 (Freshney ed.); Immobilized Cells and Enzymes, 1986 (IRL Press); Perbal, 1984, A Practical Guide to Molecular Cloning; the series, Methods in Enzymology (Academic Press, Inc.); Gene Transfer Vectors for Mammalian Cells, 15 1987 (J. H. Miller and M. P. Calos eds., Cold Spring Harbor Laboratory); Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively), all of which are incorporated by reference.

20 **IV. SEQUENCE VARIATIONS**

[0116] Numerous nucleic acid and polypeptide sequences are within the scope of the present invention, whether as target sequences or the agents used to detect those target sequences.

**SILENT VARIATIONS**

25 [0117] It will be appreciated by those skilled in the art that due to the degeneracy of the genetic code, a multitude of SARS-CoV nucleic acids sequences encoding polypeptides may be produced, some of which may bear minimal sequence homology to the nucleic acid sequences explicitly disclosed herein. For instance, inspection of the codon table (Table II) shows that codons AGA, AGG, CGA, CGC, 30 CGG, and CGU all encode the amino acid arginine. Thus, at every position in the

nucleic acids of the invention where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described above without altering the encoded polypeptide. It is understood that U in an RNA sequence corresponds to T in a DNA sequence.

5

**TABLE II**  
**Codon Table**

Amino acids			Codon				
Alanine	Ala	A	GCA	GCC	GCG	GCU	
Cysteine	Cys	C	UGC	UGU			
Aspartic acid	Asp	D	GAC	GAU			
Glutamic acid	Glu	E	GAA	GAG			
Phenylalanine	Phe	F	UUC	UUU			
Glycine	Gly	G	GGA	GGC	GGG	GGU	
Histidine	His	H	CAC	CAU			
Isoleucine	Ile	I	AUA	AUC	AUU		
Lysine	Lys	K	AAA	AAG			
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG
Methionine	Met	M	AUG				
Asparagine	Asn	N	AAC	AAU			
Proline	Pro	P	CCA	CCC	CCG	CCU	
Glutamine	Gln	Q	CAA	CAG			
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG
Threonine	Thr	T	ACA	ACC	ACG	ACU	
Valine	Val	V	GUU	GUC	GUG	GUU	
Tryptophan	Trp	W	UGG				
Tyrosine	Tyr	Y	UAC	UAU			

[0118] Such "silent variations" are one species of "conservatively modified variations", discussed below. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified by standard techniques to encode a functionally identical polypeptide. Accordingly, each silent variation of a nucleic acid, which encodes a polypeptide is implicit in any described sequence. For example, the invention provides each and every possible variation of nucleic acid sequence encoding an RdRp of a SARS-CoV that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code (e.g., as set forth in Table

II) as applied to the nucleic acid sequences encoding, e.g., an RdRp of a SARS-CoV. All such variations of every nucleic acid herein are specifically provided and described by consideration of the sequence in combination with the genetic code.

### CONSERVATIVE VARIATIONS

5 [0119] "Conservatively modified variations" or, simply, "conservative variations" of a particular nucleic acid sequence refers to those nucleic acids, which encode identical or essentially identical amino acid sequences, or, where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. One of skill will recognize that individual substitutions, deletions or additions which alter, 10 add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 4%, 2% or 1%) in an encoded sequence are "conservatively modified variations" where the alterations result in the deletion of an amino acid, addition of an amino acid, or substitution of an amino acid with a chemically similar amino acid.

15 [0120] Conservative substitution tables providing functionally similar amino acids are well known in the art. Table III sets forth six groups, which contain amino acids that are "conservative substitutions" for one another.

TABLE III  
Conservative Substitution Groups

20

1	Alanine (A)	Serine (S)	Threonine (T)
2	Aspartic acid (D)	Glutamic acid (E)	
3	Asparagine (N)	Glutamine (Q)	
4	Arginine (R)	Lysine (K)	
5	Isoleucine (I)	Leucine (L)	Methionine (M)      Valine (V)
6	Phenylalanine (F)	Tyrosine (Y)	Tryptophan (W)

25

[0121] Thus, "conservatively substituted variations" of, e.g., an RdRp of a SARS-CoV include substitutions of a small percentage, typically less than 5%, more typically less than 2% or 1%, of the amino acids of the polypeptide sequence, with a conservatively selected amino acid of the same conservative substitution group.

[0122] The addition of sequences that do not alter the encoded activity of a nucleic acid molecule, such as the addition of a non-functional sequence, is a conservative variation of the basic nucleic acid.

[0123] One of skill will appreciate that many conservative variations of the 5 nucleic acids described herein yield a functionally identical nucleic acid. For example, as discussed above, owing to the degeneracy of the genetic code, "silent substitutions" (i.e., substitutions in a nucleic acid sequence which do not result in an alteration in an encoded polypeptide) are an implied feature of *every* nucleic acid sequence, which encodes an amino acid. Similarly, "conservative amino acid substitutions," in one or a 10 few amino acids in an amino acid sequence are substituted with different amino acids with highly similar properties, are also readily identified as being highly similar to a disclosed construct. Such conservative variations of each disclosed sequence are a feature of the present invention.

## V. OLIGONUCLEOTIDE SYNTHESIS

[0124] The oligonucleotides of the invention are optionally prepared using 15 essentially any technique known in the art. In certain embodiments, for example, the oligonucleotides described herein are synthesized chemically using essentially any nucleic acid synthesis method, including, e.g., according to the solid phase phosphoramidite triester method described by Beaucage and Caruthers (1981), 20 Tetrahedron Letts. 22(20):1859-1862, which is incorporated by reference, or another synthesis technique known in the art, e.g., using an automated synthesizer, as described in Needham-VanDevanter et al. (1984) Nucleic Acids Res. 12:6159-6168, which is incorporated by reference. A wide variety of equipment is commercially available for automated oligonucleotide synthesis. Multi-nucleotide synthesis approaches (e.g., tri- 25 nucleotide synthesis, etc.) are also optionally utilized. Moreover, the primer nucleic acids described herein optionally include various modifications. In certain embodiments, for example, primers include restriction site linkers, e.g., to facilitate subsequent amplicon cloning or the like. To further illustrate, primers are also optionally modified to improve the specificity of amplification reactions as described 30 in, e.g., U.S. Pat. No. 6,001,611, entitled "MODIFIED NUCLEIC ACID AMPLIFICATION PRIMERS," issued December 14, 1999 to Will, which is

incorporated by reference. Primers and probes can also be synthesized with various other modifications as described herein or as otherwise known in the art.

[0125] Essentially any label is optionally utilized to label the nucleic acid detection reagents of the invention. In some embodiments, for example, the label 5 comprises a fluorescent dye (e.g., a rhodamine dye (e.g., R6G, R110, TAMRA, ROX, etc.), a fluorescein dye (e.g., JOE, VIC, TET, HEX, FAM, etc.), a halofluorescein dye, a cyanine dye (e.g., CY3, CY3.5, CY5, CY5.5, etc.), a BODIPY® dye (e.g., FL, 530/550, TR, TMR, etc.), an ALEXA FLUOR® dye (e.g., 488, 532, 546, 568, 594, 555, 653, 647, 660, 680, etc.), a dichlororhodamine dye, an energy transfer dye (e.g., 10 BIGDYE™ v 1 dyes, BIGDYE™ v 2 dyes, BIGDYE™ v 3 dyes, etc.), Lucifer dyes (e.g., Lucifer yellow, etc.), CASCADE BLUE®, Oregon Green, and the like. Additional examples of fluorescent dyes are provided in, e.g., Haugland, Molecular Probes Handbook of Fluorescent Probes and Research Products, Ninth Ed. (2003) and the updates thereto, which are each incorporated by reference. Fluorescent dyes are 15 generally readily available from various commercial suppliers including, e.g., Molecular Probes, Inc. (Eugene, OR), Amersham Biosciences Corp. (Piscataway, NJ), Applied Biosystems (Foster City, CA), etc. Other labels include, e.g., biotin, weakly fluorescent labels (Yin et al. (2003) Appl Environ Microbiol. 69(7):3938, Babendure et al. (2003) Anal. Biochem. 317(1):1, and Jankowiak et al. (2003) Chem Res Toxicol. 20 16(3):304), non-fluorescent labels, colorimetric labels, chemiluminescent labels (Wilson et al. (2003) Analyst. 128(5):480 and Roda et al. (2003) Luminescence 18(2):72), Raman labels, electrochemical labels, bioluminescent labels (Kitayama et al. (2003) Photochem Photobiol. 77(3):333, Arakawa et al. (2003) Anal. Biochem. 314(2):206, and Maeda (2003) J. Pharm. Biomed. Anal. 30(6):1725), and an alpha- 25 methyl-PEG labeling reagent as described in, e.g., U.S. Provisional Patent Application No. 60/428,484, filed on Nov. 22, 2002, which references are each incorporated by reference. Nucleic acid labeling is also described further below.

[0126] In addition, essentially any nucleic acid (and virtually any labeled nucleic acid, whether standard or non-standard) can be custom or standard ordered 30 from any of a variety of commercial sources, such as The Midland Certified Reagent

Company, The Great American Gene Company, ExpressGen Inc., Operon Technologies Inc., Proligo LLC, and many others.

## VI. SAMPLE PREPARATION AND NUCLEIC ACID AMPLIFICATION

[0127] Samples are generally derived or isolated from subjects, typically 5 mammalian subjects, more typically human subjects, suspected of having a SARS-CoV infection. Exemplary samples or specimens include blood, plasma, serum, feces, bronchoalveolar lavage, nasal pharyngeal swabs and tissues, urine, synovial fluid, seminal fluid, seminal plasma, prostatic fluid, vaginal fluid, cervical fluid, uterine fluid, cervical scrapings, amniotic fluid, anal scrapings, mucus, sputum, tissue, and the like.

10 10 Essentially any technique for acquiring these samples is optionally utilized including, e.g., scraping, venipuncture, swabbing, biopsy, or other techniques known in the art. Methods of storing specimens, culturing cells, isolating and preparing nucleic acids from these sources are generally known in the art and many of these are described further in the references and/or examples provided herein. A detailed description of 15 recommended facilities, practices, and protective equipment for the various laboratory biosafety levels (BSLs) may also be found in the CDC/NIH Biosafety in Microbiological and Biomedical Laboratories manual (BMBL), which is available on the world wide web at [cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm](http://cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm) as of April 20, 2004.

20 20 [0128] To illustrate, blood and urine specimens may be handled using Standard Precautions (previously Universal Precautions) in BSL-2 laboratories. Laboratory workers should wear protective equipment, including disposable gloves, laboratory coats, eye protection and a surgical mask, or face shield to provide a barrier to mucosal surface exposure. Careful attention should be given to hand hygiene after removal of 25 gloves and especially before touching the eyes or mucosal surfaces.

25 [0129] Any procedure with the potential to generate fine particulate aerosols (e.g. vortexing or sonication of specimens in an open tube) should be performed in a biological safety cabinet (BSC). The use of sealed centrifuge rotors or sample cups, if available, should be employed for centrifugation. Ideally, these rotors or cups should 30 be unloaded in a BSC.

[0130] Procedures performed outside of a BSC should be performed in a manner that minimizes the risk of exposure to an inadvertent sample release.

[0131] Work surfaces and equipment should be decontaminated after specimens are processed. Standard decontamination agents that are effective against lipid-enveloped viruses should be sufficient.

5 [0132] If the safety equipment described above is not available, administrative measures and/or additional personal protective equipment may be employed to reduce risk. This should be done in the context of a careful risk assessment by the laboratory safety officer. For example, the workflow of the laboratory may be adjusted so that a 10 minimum number of workers are present during centrifugation.

15 [0133] Consideration may be given to implementing respiratory protection for workers for use during centrifugation or other procedures with increased potential for inadvertent sample release. Acceptable methods of respiratory protection include a properly fit tested NIOSH approved filter respirator (N-95 or higher); or powered air-purifying respirators (PAPRs) equipped with high efficiency particulate air (HEPA) filters. Accurate fit testing is a key component of effective respirator use. Personnel who cannot wear fitted respirators because of facial hair or other fit-limitations should wear loose fitting hooded or helmeted PAPRs.

20 [0134] Consideration may also be given to referral of specimens to a suitably equipped reference laboratory.

25 [0135] To further illustrate, prior to analyzing the target nucleic acids described herein, those nucleic acids may be purified or isolated from samples that typically include complex mixtures of different components. Cells in collected samples are typically lysed to release the cell contents. For example, cells suspected of being infected with SARS-CoV in the particular sample can be lysed by contacting them with various enzymes, chemicals, and/or lysed by other approaches known in the art. In some embodiments, nucleic acids are analyzed directly in the cell lysate. In other embodiments, nucleic acids are further purified or extracted from cell lysates prior to detection. Essentially any nucleic acid extraction methods can be used to purify nucleic 30 acids in the samples utilized in the methods of the present invention. Exemplary

techniques that can be used to purifying nucleic acids include, e.g., affinity chromatography, hybridization to probes immobilized on solid supports, liquid-liquid extraction (e.g., phenol-chloroform extraction, etc.), precipitation (e.g., using ethanol, etc.), extraction with filter paper, extraction with micelle-forming reagents (e.g., cetyl-5 trimethyl-ammonium-bromide, etc.), binding to immobilized intercalating dyes (e.g., ethidium bromide, acridine, etc.), adsorption to silica gel or diatomic earths, adsorption to magnetic glass particles or organo silane particles under chaotropic conditions, and/or the like. Sample processing is also described in, e.g., US Pat. Nos. 5,155,018, 6,383,393, and 5,234,809, which are each incorporated by reference.

10 [0136] To further exemplify, unmodified nucleic acids can bind to a material with a silica surface. Many of these processes that are optionally adapted for use in the performing the methods of the present invention are described in the art. To illustrate, Vogelstein et al. (1979) Proc. Natl. Acad. Sci. USA 76:615-619, which is incorporated by reference, describes the purification of nucleic acids from agarose gels in the presence of sodium iodide using ground flint glass. Marko et al. (1982) Anal. Biochem. 121:382-387, which is incorporated by reference, describes the purification of nucleic acids from bacteria on glass dust in the presence of sodium perchlorate. In 15 DE-A 3734442, which is incorporated by reference, nucleic acids are isolated on glass fiber filters. The nucleic acids bound to these glass fiber filters are washed and then eluted with a methanol-containing Tris/EDTA buffer. A similar procedure is described in Jakobi et al. (1988) Anal. Biochem. 175:196-201, which is incorporated by reference. In particular, Jakobi et al. describes the selective binding of nucleic acids to 20 glass surfaces in chaotropic salt solutions and separating the nucleic acids from contaminants, such as agarose, proteins, and cell residue. To separate the glass particles from the contaminants, the particles can be centrifuged or fluids can be drawn through the glass fiber filters. In addition, the use of magnetic particles to immobilize nucleic acids after precipitation by adding salt and ethanol is described in, e.g., Alderton et al. (1992) Anal. Biochem. 201:166-169 and PCT/GB91/00212, which are 25 both incorporated by reference. In this procedure, the nucleic acids are agglutinated along with the magnetic particles. The agglutinate is separated from the original solvent by applying a magnetic field and performing one or more washing steps. After 30 at least one wash step, the nucleic acids are typically dissolved in a Tris buffer.

[0137] Magnetic particles in a porous glass matrix that is covered with a layer that includes, e.g., streptavidin can also be utilized in certain embodiments of the invention. These particles can be used, e.g., to isolate biotin-conjugated nucleic acids and proteins. Ferrimagnetic, ferromagnetic, and superparamagnetic particles are also 5 optionally utilized. Magnetic glass particles and related methods that can be adapted for using in performing the methods described herein are also described in, e.g., WO 01/37291, which is incorporated by reference.

[0138] One of the most powerful and basic technologies for deriving and detecting nucleic acids is nucleic acid amplification. In the present invention, 10 amplification of nucleic acids of interest typically precedes or is concurrent with the detection of that DNA. In addition, the oligonucleotides described herein are also optionally amplified, e.g., following chemical synthesis or the like. In some embodiments, detectable signals are amplified, e.g., using branched nucleic acid or other signal amplification formats known in the art.

[0139] Amplification methods that are optionally utilized or adapted for use with the oligonucleotides and methods described herein include, e.g., various polymerase or ligase mediated amplification methods, such as the polymerase chain reaction (PCR), reverse transcription PCR (RT-PCR), the ligase chain reaction (LCR), strand displacement amplification (SDA), nucleic acid sequence-based amplification 20 (NASBA), rolling circle amplification (RCA), and/or the like. Details regarding the use of these and other amplification methods can be found in various articles and/or any of a variety of standard texts, including, e.g., Berger, Sambrook, Ausubel, and PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press, Inc., San Diego, CA (1990) (Innis), Schweitzer et al. (2001) "Combining nucleic acid 25 amplification and detection," Curr Opin Biotechnol. 12(1):21-27, all of which are incorporated by reference. Many available biology texts also have extended discussions regarding PCR and related amplification methods. Nucleic acid amplification is also described in, e.g., Mullis et al., (1987) U.S. Patent No. 4,683,202 and Sooknanan and Malek (1995) Biotechnology 13:563, which are both incorporated 30 by reference. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369:684, which is incorporated by reference.

In certain embodiments, duplex PCR is utilized to amplify target nucleic acids. Duplex PCR amplification is described further in, e.g., Gabriel et al. (2003) "Identification of human remains by immobilized sequence-specific oligonucleotide analysis of mtDNA hypervariable regions I and II," Croat. Med. J. 44(3):293 and La et al. (2003)

5 "Development of a duplex PCR assay for detection of *Brachyspira hyodysenteriae* and *Brachyspira pilosicoli* in pig feces," J. Clin. Microbiol. 41(7):3372, which are both incorporated by reference. Optionally, labeled primers (e.g., biotinylated primers, Scorpion primers, etc.) are utilized to amplify nucleic acids in a sample, e.g., to facilitate the detection of amplicons and the like. Scorpion primers are also described  
10 in, e.g., Whitcombe et al. (1999) "Detection of PCR products using self-probing amplicons and fluorescence" Nat Biotechnol. 17(8):804-807, which is incorporated by reference. Labeling is described further herein.

[0140] Amplicons are optionally recovered and purified from other reaction components by any of a number of methods well known in the art, including  
15 electrophoresis, chromatography, precipitation, dialysis, filtration, and/or centrifugation. Aspects of nucleic acid purification are described in, e.g., Douglas et al., DNA Chromatography, Wiley, John & Sons, Inc. (2002), and Schott, Affinity Chromatography: Template Chromatography of Nucleic Acids and Proteins, Chromatographic Science Series, #27, Marcel Dekker (1984), all of which are  
20 incorporated by reference. In certain embodiments, amplicons are not purified prior to detection. The detection of amplicons is described further below.

#### NESTED PCR

[0141] The present invention provides methods for detecting SARS-CoV in samples and typing the coronavirus. An exemplary method is as follows:

25 (a) Nested PCR using two sets of amplification primers. The target DNA sequence of one set of primer (termed "inner" primers) is located within the target sequence of the second set of primers (termed "outer" primers). In practice, a standard PCR reaction is first run with the patient sample using the "outer primers". Then a second PCR reaction is run with the "inner primers" using the product of the  
30 first reaction as the amplification target. This procedure increases the sensitivity of the assay by reamplifying the product of the first reaction in a second reaction. The

specificity of the assay is increased because the inner primers amplify only if the first PCR reaction yielded a specific product. A schematic depiction of a nested PCR procedure is shown in Figure 3.

(b) Treating the sample with consensus SARS virus primer nucleic acids, an agent for polymerization, and deoxynucleoside 5'-triphosphates under hybridizing conditions, in which consensus SARS virus primers are mixtures of oligonucleotides that comprise at least one pair of primers sufficiently complementary to the target region of the SARS-CoV genome to hybridize therewith such that extension products or amplicons are synthesized from at least one member of the pair.

10 When separated from complementary strands, the amplicons can serve as templates for synthesis of extension products or amplicons of the other member of the pair of primer nucleic acids;

(c) Treating the sample under denaturing conditions to separate the primer extension products or amplicons from their templates to provide single-stranded molecules;

15 (d) Treating the products of step (b) with the consensus primers of the step (a) under conditions such that additional primer extension products are synthesized using the single-stranded molecules produced in step (c) as a template;

(e) Repeating steps (b) - (d) to synthesize detectable amounts of target SARS-CoV amplicons to providing amplified SARS-CoV sequences;

20 (f) Determining if amplification has occurred by treating the reaction mixture prepared in step (d) under hybridizing conditions with a generic probe and determining if hybridization and amplification has occurred; and,

(g) Hybridizing a type-specific DNA probe to the amplified DNA and determining if hybridization has occurred.

25 [0142] Amplification of nucleic acids by the polymerase chain reaction is also described in U.S. Pat. Nos. 4,683,202 and 4,683,195, which are both incorporated by reference. In general, PCR amplification of nucleic acids involves repeated cycles of heat-denaturing the nucleic acids, annealing oligonucleotide primers to sequences that flank the nucleic acid segment to be amplified, and extending the annealed primers with a polymerase. The primers hybridize to opposite strands of the target sequence and are oriented such that nucleic acid synthesis by the polymerase proceeds across the region

between the primers, effectively doubling the amount of the target nucleic acid synthesized in the previous cycle. This results in the exponential accumulation of the target nucleic acids.

[0143] The choice of primers for use in PCR typically determines the 5 specificity of the amplification reaction. In the amplification steps of the method of the present invention, "consensus" primers are used that will amplify SARS-CoV sequences present in a sample, regardless of type. The consensus primers of the invention can include degenerate primers, mixtures of the oligonucleotides synthesized so that any one several nucleotides can be incorporated into a primer are sufficiently 10 complementary to all types of SARS-CoV to amplify a nucleic acid sequence of any SARS-CoV present in the sample.

[0144] The present invention may also lead to the discovery of previously unknown strains or types of SARS-CoVs.

#### 5'-NUCLEASE REACTIONS

[0145] The oligonucleotides of the invention can also be utilized as 5'-nuclease probes for the real-time detection of SARS-CoV in 5'-nuclease PCR reactions.

[0146] 5'-nuclease reactions typically use 5'-nuclease probes to detect specific PCR products by employing the 5'-->3' exonuclease activity of, e.g., *Taq* DNA polymerase. In certain embodiments, 5'-nuclease probes are blocked to prohibit 20 extension at their 3' ends. 5'-nuclease probes typically include site-specific sequences and are labeled with a fluorescent reporter dye and a fluorescent quencher dye. Typically, 5'-nuclease probes hybridize to target sequences downstream from the 3' ends of primer sequences used for the PCR reactions. During PCR the 5'-nuclease probes hybridizes to its complementary single strand nucleic sequence within the target 25 region. When amplification occurs 5'-nuclease probes are degraded due to the 5'-->3' exonuclease activity of, e.g., *Taq* DNA polymerase, thereby separating the quencher from the reporter during extension. Due to the release of the quenching effect on the reporter, the detectable fluorescence intensity of the reporter dye increases. During the amplification process this light emission generally increases exponentially. Light 30 emission from the reporter dye is typically monitored spectrophotometrically. Because

increase of the fluorescence intensity of the reporter dye is only achieved when probe hybridization and amplification of the target sequence has occurred, 5'-nuclease assays offer sensitive methods (viral nucleic acids at a concentration of only a few cfu/ml can be detected) to determine the presence or absence of specific sequences. Therefore, 5 this technique is particularly useful in diagnostic applications, such as the screening of samples for the detection of SARS-CoV in clinical samples.

[0147] Exemplary probes useful for detection of the SARS-CoV in samples in 5'-nuclease reactions include, e.g., oligonucleotides having sequences corresponding to SEQ ID NOS: 17 and 21. To illustrate, a 5'-nuclease probe can incorporate a FAM 10 fluorophore at the 5' end and the TAMRA quencher at the 3' end. To further illustrate, a 5'-nuclease probe with a sequence corresponding to SEQ ID NO: 21 is optionally used in real-time PCR reactions with primer nucleic acids having sequences corresponding to SEQ ID NO: 11 and SEQ ID NO: 20 in certain embodiments of the invention.

15 [0148] Example conditions optionally used in 5'-nuclease reactions to detect SARS-CoV in samples are shown below. Reaction Conditions II are generally more useful if the target nucleic acid is expected to be present in a relatively high proportion. Optimization of reaction conditions is considered within the skill of the ordinary practitioner of the art.

#### REACTION CONDITIONS I

2.5 $\mu$ l 10X TaqMan® A Buffer	1X
3.0 $\mu$ l 25mM MgCl <sub>2</sub>	3mM
.75 $\mu$ l 10mM dATP	300 $\mu$ M
.75 $\mu$ l 10mM dCTP	300 $\mu$ M
.75 $\mu$ l 10mM dGTP	300 $\mu$ M

.75 $\mu$ l 20mM dUTP	600 $\mu$ M
.075 $\mu$ l 100 $\mu$ M forward primer	300 nM
.225 $\mu$ l 100 $\mu$ M reverse primer	900 nM
.25 $\mu$ l 10 $\mu$ M probe (SEQ ID NO: 17 or 21)	100nM
.125 $\mu$ l AmpliTaq Gold® (5 U/ $\mu$ l)	.025 U/ $\mu$ l
.25 $\mu$ l AmpErase® UNG (1 U/ $\mu$ l)	.01 U/ $\mu$ l
<b>13.575 <math>\mu</math>l Nuclease-free water</b>	
<b>2.0 <math>\mu</math>l DNA standard/sample</b>	

**total volume = 25  $\mu$ l**

#### **CYCLING CONDITIONS**

**50°C for 2 min**  
**60°C for 30 min**  
**95°C for 5 min**  
**40 cycles of 95°C 15 and 60°C  
for 1 min**  
**25°C for 2 min**

#### **REACTION CONDITIONS II**

<b>5 <math>\mu</math>l 10X TaqMan® Buffer A</b>	<b>1X</b>
<b>11 <math>\mu</math>l 25mM MgCl<sub>2</sub></b>	<b>5.5mM</b>

11 $\mu$ l 25mM MgCl <sub>2</sub>	5.5mM
1 $\mu$ l 10mM dATP	200 $\mu$ M
1 $\mu$ l 10mM dCTP	200 $\mu$ M
1 $\mu$ l 10mM dGTP	200 $\mu$ M
1 $\mu$ l 20mM dUTP	400 $\mu$ M
1 $\mu$ l 10 $\mu$ M forward primer	200 nM
1 $\mu$ l 10 $\mu$ M reverse primer	200 nM
1 $\mu$ l 10 $\mu$ M probe (SEQ ID NO: 17 or 21)	100nM
.5 AmpliTaq Gold® (5 U/ $\mu$ l)	.05 U/ $\mu$ l
.5 $\mu$ l AmpErase® UNG (1 U/ $\mu$ l)	.01 U/ $\mu$ l
1.25 $\mu$ l 2% glycerin	0.05%
23.08 $\mu$ l RNase-free water	
1.67 $\mu$ l DNA standard/sample	

total volume = 50  $\mu$ l

#### CYCLING CONDITIONS

50°C for 2 min

95°C for 5 min

40 cycles of 95°C for 15  
sec and 60°C for 1 min

25°C for 2 min

[0149] An exhaustive manual for conducting RT-PCR using 5'-nuclease reactions is available from Applied Biosystems, "TaqMan® One-Step RT-PCR Master Mix Reagents Kit Protocol", printed 4/2002, as part no. 4310299 Rev.C., hereby 5 incorporated in its entirety by reference. The publication was available on the world wide web at [ucl.ac.uk/wibr/2/services/reldocs/1steppcr.pdf](http://ucl.ac.uk/wibr/2/services/reldocs/1steppcr.pdf) as of April 20, 2004.

## VII. OLIGONUCLEOTIDE ARRAYS

[0150] In certain embodiments of the invention, the oligonucleotides described herein are covalently or noncovalently attached to solid supports which are then 10 contacted with samples comprising amplified and labeled nucleic acid from a subject. In other embodiments, the oligonucleotides of the invention are provided free in solution. Essentially any substrate material is optionally adapted for use in these aspects of the invention. In certain embodiments, for example, substrates are fabricated 15 from silicon, glass, or polymeric materials (e.g., glass or polymeric microscope slides, silicon wafers, etc.). Suitable glass or polymeric substrates, including microscope slides, are available from various commercial suppliers, such as Fisher Scientific (Pittsburgh, PA) or the like. In some embodiments, solid supports utilized in the invention are membranes. Suitable membrane materials are optionally selected from, e.g. polyaramide membranes, polycarbonate membranes, porous plastic matrix 20 membranes (e.g., POREX® Porous Plastic, etc.), porous metal matrix membranes, polyethylene membranes, poly(vinylidene difluoride) membranes, polyamide membranes, nylon membranes, ceramic membranes, polyester membranes, polytetrafluoroethylene (TEFLON®) membranes, woven mesh membranes, microfiltration membranes, nanofiltration membranes, ultrafiltration membranes, 25 dialysis membranes, composite membranes, hydrophilic membranes, hydrophobic

membranes, polymer-based membranes, a non-polymer-based membranes, powdered activated carbon membranes, polypropylene membranes, glass fiber membranes, glass membranes, nitrocellulose membranes, cellulose membranes, cellulose nitrate membranes, cellulose acetate membranes, polysulfone membranes, polyethersulfone membranes, polyolefin membranes, or the like. Many of these membranous materials are widely available from various commercial suppliers, such as, P.J. Cobert Associates, Inc. (St. Louis, MO), Millipore Corporation (Bedford, MA), or the like. Other exemplary solid supports that are optionally utilized include, e.g., ceramics, metals, resins, gels, plates, beads, microbeads (e.g., magnetic microbeads, etc.), tubes (e.g., microtubes, etc.), whiskers, fibers, combs, single crystals, and self-assembling monolayers.

[0151] The oligonucleotides of the invention are directly or indirectly (e.g., via linkers, such as bovine serum albumin (BSA) or the like) attached to the supports, e.g., by any available chemical or physical method. A wide variety of linking chemistries are available for linking molecules to a wide variety of solid supports. More specifically, nucleic acids may be attached to the solid support by covalent binding such as by conjugation with a coupling agent or by non-covalent binding such as electrostatic interactions, hydrogen bonds or antibody-antigen coupling, or by combinations thereof. Typical coupling agents include biotin/avidin, biotin/streptavidin, *Staphylococcus aureus* protein A/IgG antibody F<sub>c</sub> fragment, and streptavidin/protein A chimeras (Sano et al. (1991) *Bio/Technology* 9:1378, which is incorporated by reference), or derivatives or combinations of these agents. Nucleic acids may be attached to the solid support by a photocleavable bond, an electrostatic bond, a disulfide bond, a peptide bond, a diester bond or a combination of these bonds. Nucleic acids are also optionally attached to solid supports by a selectively releasable bond such as 4,4'-dimethoxytrityl or its derivative. Derivatives which have been found to be useful include 3 or 4 [bis-(4-methoxyphenyl)]-methyl-benzoic acid, N-succinimidyl-3 or 4 [bis-(4-methoxyphenyl)]-methyl-benzoic acid, N-succinimidyl-3 or 4 [bis-(4-methoxyphenyl)]-hydroxymethyl-benzoic acid, N-succinimidyl-3 or 4 [bis-(4-methoxyphenyl)]-chloromethyl-benzoic acid, and salts of these acids.

[0152] As referred to above, oligonucleotides are optionally attached to solid supports via linkers between the nucleic acids and the solid support. Useful linkers include a coupling agent, as described above for binding to other or additional coupling partners, or to render the attachment to the solid support cleavable.

5 [0153] Cleavable attachments can be created by attaching cleavable chemical moieties between the oligonucleotides and the solid support including, e.g., an oligopeptide, oligonucleotide, oligopolyamide, oligoacrylamide, oligoethylene glycerol, alkyl chains of between about 6 to 20 carbon atoms, and combinations thereof. These moieties may be cleaved with, e.g., added chemical agents, electromagnetic radiation, 10 or enzymes. Exemplary attachments cleavable by enzymes include peptide bonds which can be cleaved by proteases, and phosphodiester bonds which can be cleaved by nucleases.

15 [0154] Chemical agents such as  $\beta$ -mercaptoethanol, dithiothreitol (DTT) and other reducing agents cleave disulfide bonds. Other agents which may be useful include oxidizing agents, hydrating agents and other selectively active compounds. Electromagnetic radiation such as ultraviolet, infrared and visible light cleave photocleavable bonds. Attachments may also be reversible, e.g., using heat or enzymatic treatment, or reversible chemical or magnetic attachments. Release and reattachment can be performed using, e.g., magnetic or electrical fields.

20 [0155] Array based hybridization is particularly suitable for detecting SARS-CoV nucleic acids, as it can be used to detect the presence of many amplicons simultaneously. A number of array systems have been described and can be adapted for use with the present invention, including those available from commercial suppliers such as Affymetrix, Inc. (Santa Clara, CA, USA) and the like. Aspects of array 25 construction and use are also described in, e.g., Sapolisky et al. (1999) "High-throughput polymorphism screening and genotyping with high-density oligonucleotide arrays." Genetic Analysis: Biomolecular Engineering 14:187-192; Lockhart (1998) "Mutant yeast on drugs" Nature Medicine 4:1235-1236; Fodor (1997) "Genes, Chips and the Human Genome." FASEB Journal 11:A879; Fodor (1997) "Massively Parallel 30 Genomics" Science 277: 393-395; and Chee et al. (1996) "Accessing Genetic

Information with High-Density DNA Arrays" Science 274:610-614, all of which are incorporated by reference.

### VIII. NUCLEIC ACID HYBRIDIZATION

[0156] Hybridization of oligonucleotides to their target SARS-CoV nucleic

5 acids can be accomplished by choosing the appropriate hybridization conditions. The stability of the probe:target nucleic acid hybrid is typically selected to be compatible with the assay and washing conditions so that stable, detectable hybrids form only between the probes and target SARS-CoV nucleic acids. Manipulation of one or more of the different assay parameters determines the exact sensitivity and specificity of a  
10 particular hybridization assay.

[0157] More specifically, hybridization between complementary bases of DNA, RNA, PNA, or combinations of DNA, RNA and PNA, occurs under a wide variety of conditions that vary in temperature, salt concentration, electrostatic strength, buffer composition, and the like. Examples of these conditions and methods for applying

15 them are described in, e.g., Tijssen (1993), *supra*, and Hames and Higgins, *supra*. Hybridization generally takes place between about 0°C and about 70°C, for periods of from about one minute to about one hour, depending on the nature of the sequence to be hybridized and its length. However, it is recognized that hybridizations can occur in seconds or hours, depending on the conditions of the reaction. To illustrate, typical  
20 hybridization conditions for a mixture of two 20-mers is to bring the mixture to 68°C, followed by cooling to room temperature (22°C) for five minutes or at very low temperatures such as 2°C in 2 microliters. Hybridization between nucleic acids may be facilitated using buffers such as Tris-EDTA (TE), Tris-HCl and HEPES, salt solutions (e.g. NaCl, KC1, CaCl<sub>2</sub>), or other aqueous solutions, reagents and chemicals.

25 Examples of these reagents include single-stranded binding proteins such as Rec A protein, T4 gene 32 protein, *E. coli* single-stranded binding protein and major or minor nucleic acid groove binding proteins. Other examples of such reagents and chemicals include divalent ions, polyvalent ions and intercalating substances such as ethidium bromide, actinomycin D, psoralen, and angelicin.

**IX. DETECTION AND OLIGONUCLEOTIDE VARIATIONS**

[0158] As referred to above, amplified target SARS-CoV nucleic acid in the

samples utilized in the methods of the invention is optionally labeled to permit

detection of oligonucleotide-target hybridization duplexes. In general, a label can be

5 any moiety that can be attached, e.g., to a primer utilized for amplification and provide a detectable signal (e.g., a quantifiable signal). Labels may be attached to a primer directly or indirectly by a variety of techniques known in the art. Depending on the type of label used, the label can be attached to a terminal (5' or 3' end of the primer) or a non-terminal nucleotide, and can be attached indirectly through linkers or spacer arms

10 of various sizes and compositions. Using commercially available phosphoramidite reagents, one can produce oligomers containing functional groups (e.g., thiols or primary amines) at either the 5' or 3' terminus via an appropriately protected phosphoramidite, and can label such oligonucleotides using protocols described in, for example, PCR Protocols: A Guide to Methods and Applications (Innis et al, eds.

15 Academic Press, Inc. (1990)). In one embodiment, the label consists of a biotin molecule covalently bound to the primer at the 5' end. The term "biotinylated primer" refers to a primer with one or more biotin molecules bound either directly to the primer or indirectly through intervening linker molecules.

[0159] To further illustrate, detection of oligonucleotide-target hybridization duplexes is optionally by a chemiluminescent assay using a luminol-based reagent as described in, e.g., Whitehead, et al. (1983) Nature 30(5):158, which is incorporated by reference, and available commercially. Following hybridization of the oligonucleotide with the labeled target DNA, the biotin molecule attached to the target DNA is conjugated, e.g., to streptavidin-horseradish peroxidase (SA-HRP). Alternatively, the target DNA can be labeled with horseradish peroxidase directly, thereby eliminating the separate conjugation step. In either case, subsequent oxidation of luminol by the horseradish peroxidase enzyme results in the emission of photons, which is then detected, e.g., on standard autoradiography film. The intensity of the signal is a function of DNA quantity. A series of DNA standards containing known amounts of DNA are typically assayed along with one or more unknown samples. The signal intensities of the known DNA standards allow an empirical determination of the functional relationship between signal intensity and DNA quantity, which enables the

quantitation of the unknown samples. Many other methods of detection are also optionally utilized to perform the methods of the invention and are referred to in the references cited herein and/or generally known in the art.

[0160] Any available method for detecting SARS-CoV amplicons can be used 5 in the present invention. Common approaches include real time amplification detection with molecular beacons or 5'-nuclease probes, detection of intercalating dyes, detection of labels incorporated into the amplification probes or the amplified nucleic acids themselves, e.g., following electrophoretic separation of the amplification products from unincorporated label), hybridization based assays (e.g., array based assays) and/or 10 detection of secondary reagents that bind to the nucleic acids.

[0161] To further illustrate, a molecular beacon or a 5'-nuclease probe is 15 optionally designed to include a oligonucleotide of the invention (i.e., is selected from SEQ ID NOS: 1-12 and 15-24) or complements thereto), which molecular beacon or 5'-nuclease probe can be used to detect SARS-CoV amplicons. Molecular beacons or 5'-nuclease probes are described further below. Details on these general approaches are 20 found in the references cited herein, e.g., Sambrook and Ausubel. Additional labeling strategies for labeling nucleic acids and corresponding detection strategies can be found, e.g., in Haugland (2003) Handbook of Fluorescent Probes and Research Chemicals Ninth Edition by Molecular Probes, Inc. (Eugene, OR), which is incorporated by reference.

[0162] Molecular beacons (MBs) are oligonucleotides designed for real time 25 detection and quantification of target nucleic acids (e.g., target SARS-CoV amplicons). The 5' and 3' termini of MBs collectively comprise a pair of moieties which confers the detectable properties of the MB. One of the termini is attached to a fluorophore and the other is attached to a quencher molecule capable of quenching a fluorescent 30 emission of the fluorophore. For example, one example fluorophore-quencher pair can use a fluorophore such as EDANS or fluorescein, e.g., on the 5'-end and a quencher such as Dabcyl, e.g., on the 3'-end. When the MB is present free in solution, i.e., not hybridized to a second nucleic acid, the stem of the MB is stabilized by complementary base pairing. This self-complementary pairing results in a "hairpin loop" structure for the MB in which the fluorophore and the quenching moieties are proximal to one

another. In this confirmation, the fluorescent moiety is quenched by the fluorophore. The loop of the molecular beacon typically comprises an oligonucleotide described herein (i.e., is selected from SEQ ID NOS: 1-12 and 15-24 or complements thereto) and is accordingly complementary to a sequence to be detected in the target SARS-CoV 5 nucleic acid, such that hybridization of the loop to its complementary sequence in the target forces disassociation of the stem, thereby distancing the fluorophore and quencher from each other. This results in unquenching of the fluorophore, causing an increase in fluorescence of the MB.

[0163] Details regarding standard methods of making and using MBs are well 10 established in the literature and MBs are available from a number of commercial reagent sources. Further details regarding methods of MB manufacture and use are found, e.g., in Leone et al. (1995) "Molecular beacon probes combined with amplification by NASBA enable homogenous real-time detection of RNA," Nucleic Acids Res. 26:2150-2155; Hsuih et al. (1997) "Novel, ligation-dependent PCR assay 15 for detection of hepatitis C in serum" J Clin Microbiol 34:501-507; Kostrikis et al. (1998) "Molecular beacons: spectral genotyping of human alleles" Science 279:1228-1229; Sokol et al. (1998) "Real time detection of DNA:RNA hybridization in living 20 cells" Proc. Natl. Acad. Sci. U.S.A. 95:11538-11543; Tyagi et al. (1998) "Multicolor molecular beacons for allele discrimination" Nature Biotechnology 16:49-53; Fang et al. (1999) "Designing a novel molecular beacon for surface-immobilized DNA 25 hybridization studies" J. Am. Chem. Soc. 121:2921-2922; and Marras et al. (1999) "Multiplex detection of single-nucleotide variation using molecular beacons" Genet. Anal. Biomol. Eng. 14:151-156, all of which are incorporated by reference. Aspects of MB construction and use are also found in patent literature, such as U.S. Pat. No. 30 5,925,517 (July 20, 1999) to Tyagi et al. entitled "Detectably labeled dual conformation oligonucleotide probes, assays and kits;" U.S. Pat. No. 6,150,097 to Tyagi et al (November 21, 2000) entitled "Nucleic acid detection probes having non-FRET fluorescence quenching and kits and assays including such probes" and U.S. Pat. No. 6,037,130 to Tyagi et al (March 14, 2000), entitled "Wavelength-shifting probes and primers and their use in assays and kits," all of which are incorporated by reference.

[0164] MB components (e.g., oligos, including those labeled with fluorophores or quenchers) can be synthesized using conventional methods. Some of these methods are described further above. For example, oligonucleotides or peptide nucleic acids (PNAs) can be synthesized on commercially available automated oligonucleotide/PNA synthesis machines using standard methods. Labels can be attached to the oligonucleotides or PNAs either during automated synthesis or by post-synthetic reactions which have been described before *see, e.g.*, Tyagi and Kramer (1996), *supra*. Aspects relating to the synthesis of functionalized oligonucleotides can also be found in Nelson, et al. (1989) "Bifunctional Oligonucleotide Probes Synthesized Using A Novel CPG Support Are Able To Detect Single Base Pair Mutations" Nucleic Acids Res. 17:7187-7194, which is incorporated by reference. Labels/ quenchers can be introduced to the oligonucleotides or PNAs, e.g., by using a controlled-pore glass column to introduce, e.g., the quencher (e.g., a 4-dimethylaminoazobenzene-4'-sulfonyl moiety (DABSYL). For example, the quencher can be added at the 3' end of oligonucleotides during automated synthesis; a succinimidyl ester of 4-(4'-dimethylaminophenylazo)benzoic acid (DABCYL) can be used when the site of attachment is a primary amino group; and 4-dimethylaminophenylazophenyl-4'-maleimide (DABMI) can be used when the site of attachment is a sulphhydryl group. Similarly, fluorescein can be introduced in the oligonucleotides, either using a fluorescein phosphoramidite that replaces a nucleoside with fluorescein, or by using a fluorescein dT phosphoramidite that introduces a fluorescein moiety at a thymidine ring via a linker. To link a fluorescein moiety to a terminal location, iodoacetoamidofluorescein can be coupled to a sulphhydryl group. Tetrachlorofluorescein (TET) can be introduced during automated synthesis using a 5'-tetrachloro-fluorescein phosphoramidite. Other reactive fluorophore derivatives and their respective sites of attachment include the succinimidyl ester of 5-carboxyrhodamine-6G (RHD) coupled to an amino group; an iodoacetamide of tetramethylrhodamine coupled to a sulphhydryl group; an isothiocyanate of tetramethylrhodamine coupled to an amino group; or a sulfonylchloride of Texas red coupled to a sulphhydryl group. During the synthesis of these labeled components, conjugated oligonucleotides or PNAs can be purified, if desired, e.g., by high pressure liquid chromatography or other methods.

[0165] A variety of commercial suppliers produce standard and custom molecular beacons, including Cruachem ([cruachem.com](http://cruachem.com)), Oswel Research Products Ltd. (UK; [oswel.com](http://oswel.com)), Research Genetics (a division of Invitrogen, Huntsville AL ([resgen.com](http://resgen.com))), the Midland Certified Reagent Company (Midland, TX [mcrc.com](http://mcrc.com)) and

5 Gorilla Genomics, LLC (Alameda, CA). A variety of kits, which utilize molecular beacons are also commercially available, such as the Sentinel™ Molecular Beacon Allelic Discrimination Kits from Stratagene (La Jolla, CA) and various kits from Eurogentec SA (Belgium, [eurogentec.com](http://eurogentec.com)) and Isogen Bioscience BV (The Netherlands, [isogen.com](http://isogen.com)).

10 [0166] In certain embodiments, as also referred to above, a real time PCR assay system that includes one or more 5'-nuclease probes is used for detecting amplified SARS-CoV nucleic acids. These systems operate by using the endogenous nuclease activity of certain polymerases to cleave a quencher or label free from an oligonucleotide of the invention that comprises the quencher and label, resulting in unquenching of the label. The polymerase only cleaves the quencher or label upon initiation of replication, i.e., when the oligonucleotide is bound to the template and the polymerase extends the primer. Thus, an appropriately labeled oligonucleotide and polymerase comprising the appropriate nuclease activity can be used to detect a SARS-CoV nucleic acid of interest. Real time PCR product analysis by, e.g., FRET or the like

15 (and related real time reverse-transcription PCR) provides a well-known technique for real time PCR monitoring that has been used in a variety of contexts, which can be adapted for use with the probes and methods described herein (see, Laurendeau et al. (1999) "TaqMan PCR-based gene dosage assay for predictive testing in individuals from a cancer family with INK4 locus haploinsufficiency" *Clin Chem* 45(7):982-6;

20 Laurendeau et al. (1999) "Quantitation of MYC gene expression in sporadic breast tumors with a real-time reverse transcription-PCR assay" *Clin Chem* 59(12):2759-65; and Kreuzer et al. (1999) "LightCycler technology for the quantitation of bcr/abl fusion transcripts" *Cancer Research* 59(13):3171-4, all of which are incorporated by reference). To further illustrate, exemplary 5'-nuclease probes of the invention include

25 5'-(6-FAM)-AGCTAACGAGTGTGCGCAAGTATTAAGTGAGATG-(TAMRA) and 5'-(6-FAM)-AGAGCCATGCCTAACAT(NFQ)-3' (SEQ ID NO: 27) and 5'-(6-FAM)-AGAGCCATGCCTAACAT(NFQ)-3' (SEQ ID NO: 28).

## X. SYSTEMS

[0167] The invention also provides a system for detecting SARS-CoV in a sample. The system includes one or more nucleic acid detection reagents as described herein (e.g., probe nucleic acids, sequence specific antibodies, etc.). In certain 5 embodiments, the nucleic acid detection reagents are arrayed on a solid support, whereas in others, they are provided in one or more containers, e.g., for assays performed in solution. The system also includes at least one detector (e.g., a spectrometer, etc.) that detects binding between nucleic acids and/or amplicons thereof from the sample and the nucleic acid detection reagent. Other detectors are described 10 further below. In addition, the system also includes at least one controller operably connected to the detector. The controller includes one or more instructions sets that correlate the binding detected by the detector with a presence of SARS-CoV in the sample.

[0168] In some embodiments, at least one container includes the nucleic acid 15 detection reagent. In these embodiments, the system optionally further includes at least one thermal modulator operably connected to the container to modulate temperature in the container, and/or at least one fluid transfer component (e.g., an automated pipettor, etc.) that transfers fluid to and/or from the container, e.g., for performing one or more nucleic acid amplification techniques in the container, etc.

[0169] Exemplary commercially available systems that are optionally utilized to 20 detect SARS-CoV nucleic acids using the nucleic acid detection reagents described herein (e.g., oligonucleotides comprising sequences selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24 or complements thereto, sequence specific antibodies, etc.) include, e.g., a LightCycler™ system (e.g., LightCycler™ RNA Master SYBR Green Detection) or a COBAS AMPLICOR® Analyzer, which are available from Roche Diagnostics Corporation (Indianapolis, IN), a LUMINEX 100™ system, which is available from the Luminex Corporation (Austin, TX), an ABI PRISM® Sequence Detection System, which is available from Applied Biosystems (Foster City, CA), and the like.

[0170] The invention further provides a computer or computer readable 25 medium that includes a data set that comprises a plurality of character strings that

correspond to a plurality of sequences that correspond to one or more of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. Typically, the computer or computer readable medium further includes an automatic synthesizer coupled to an output of the computer or computer readable medium. The automatic synthesizer accepts instructions from the computer or computer readable medium, which instructions direct synthesis of, e.g., one or more probe nucleic acids that correspond to one or more character strings in the data set. Exemplary systems and system components are 5 described further below.

10

[0171] Detectors are structured to detect detectable signals produced, e.g., in or proximal to another component of the system (e.g., in container, on a solid support, etc.). Suitable signal detectors that are optionally utilized, or adapted for use, in these systems detect, e.g., fluorescence, phosphorescence, radioactivity, absorbance, 15 refractive index, luminescence, or the like. Detectors optionally monitor one or a plurality of signals from upstream and/or downstream of the performance of, e.g., a given assay step. For example, the detector optionally monitors a plurality of optical signals, which correspond in position to "real time" results. Example detectors or sensors include photomultiplier tubes, CCD arrays, optical sensors, temperature 20 sensors, pressure sensors, pH sensors, conductivity sensors, scanning detectors, or the like. Each of these as well as other types of sensors is optionally readily incorporated into the systems described herein. Optionally, the systems of the present invention include multiple detectors.

[0172] More specific exemplary detectors that are optionally utilized in these 25 systems include, e.g., a resonance light scattering detector, an emission spectroscope, a fluorescence spectroscope, a phosphorescence spectroscope, a luminescence spectroscope, a spectrophotometer, a photometer, and the like. Various synthetic components are also utilized, or adapted for, use in the systems of the invention including, e.g., automated nucleic acid synthesizers, e.g., for synthesizing the 30 oligonucleotides described herein. Detectors and synthetic components that are optionally included in the systems of the invention are described further in, e.g., Skoog

et al., Principles of Instrumental Analysis, 5<sup>th</sup> Ed., Harcourt Brace College Publishers (1998) and Currell, Analytical Instrumentation: Performance Characteristics and Quality, John Wiley & Sons, Inc. (2000), both of which are incorporated by reference.

[0173] The systems of the invention also typically include controllers that are operably connected to one or more components (e.g., detectors, synthetic components, thermal modulator, fluid transfer components, etc.) of the system to control operation of the components. More specifically, controllers are generally included either as separate or integral system components that are utilized, e.g., to receive data from detectors, to effect and/or regulate temperature in the containers, to effect and/or regulate fluid flow to or from selected containers, or the like. Controllers and/or other system components is/are optionally coupled to an appropriately programmed processor, computer, digital device, or other information appliance (e.g., including an analog to digital or digital to analog converter as needed), which functions to instruct the operation of these instruments in accordance with preprogrammed or user input instructions, receive data and information from these instruments, and interpret, manipulate and report this information to the user. Suitable controllers are generally known in the art and are available from various commercial sources.

[0174] Any controller or computer optionally includes a monitor, which is often a cathode ray tube ("CRT") display, a flat panel display (e.g., active matrix liquid crystal display, liquid crystal display, etc.), or others. Computer circuitry is often placed in a box, which includes numerous integrated circuit chips, such as a microprocessor, memory, interface circuits, and others. The box also optionally includes a hard disk drive, a floppy disk drive, a high capacity removable drive such as a writeable CD-ROM, and other common peripheral elements. Inputting devices such as a keyboard or mouse optionally provide for input from a user. These components are illustrated further below.

[0175] The computer typically includes appropriate software for receiving user instructions, either in the form of user input into a set of parameter fields, e.g., in a GUI, or in the form of preprogrammed instructions, e.g., preprogrammed for a variety of different specific operations. The software then converts these instructions to appropriate language for instructing the operation of one or more controllers to carry

out the desired operation. The computer then receives the data from, e.g., sensors/detectors included within the system, and interprets the data, either provides it in a user understood format, or uses that data to initiate further controller instructions, in accordance with the programming, e.g., such as controlling fluid flow regulators in 5 response to fluid weight data received from weight scales or the like.

[0176] The computer can be, e.g., a PC (Intel x86 or Pentium chip-compatible DOS™, OS2™, WINDOWS™, WINDOWS NT™, WINDOWS95™, WINDOWS98™, WINDOWS2000™, WINDOWS XPTM, LINUX-based machine, a MACINTOSH™, Power PC, or a UNIX-based (e.g., SUN™ work station) machine) or 10 other common commercially available computer, which is known to one of skill in the art. Standard desktop applications such as word processing software (e.g., Microsoft Word™ or Corel WordPerfect™) and database software (e.g., spreadsheet software such as Microsoft Excel™, Corel Quattro Pro™, or database programs such as Microsoft Access™ or Paradox™) can be adapted to the present invention. Software 15 for performing, e.g., controlling temperature modulators and fluid:flow regulators is optionally constructed by one of skill using a standard programming language such as Visual basic, Fortran, Basic, Java, or the like.

[0177] Figures 4 and 5 are schematics showing representative example systems that include logic devices in which various aspects of the present invention may be 20 embodied. As will be understood by practitioners in the art from the teachings provided herein, the invention is optionally implemented in hardware and/or software. In some embodiments, different aspects of the invention are implemented in either client-side logic or server-side logic. As will be understood in the art, the invention or components thereof may be embodied in a media program component (e.g., a fixed 25 media component) containing logic instructions and/or data that, when loaded into an appropriately configured computing device, cause that device to perform according to the invention. As will also be understood in the art, a fixed media containing logic instructions may be delivered to a viewer on a fixed media for physically loading into a viewer's computer or a fixed media containing logic instructions may reside on a 30 remote server that a viewer accesses through a communication medium in order to download a program component.

[0178] In particular, Figure 4 schematically illustrate computer 400 to which detector 402 and fluid transfer component 404 are operably connected. Optionally, detector 402 and/or fluid transfer component 404 is operably connected to computer 400 via a server (not shown in Figure 4). During operation, fluid transfer component 5 404 typically transfers fluids, such as sample aliquots comprising labeled SARS-CoV amplicons to nucleic acid detection reagent array 406, e.g., comprising oligonucleotides, sequence specific antibodies, etc., as described herein, arrayed thereon. Thereafter, detector 402 typically detects detectable signals (e.g., fluorescent emissions, etc.) produced by labeled amplicons that hybridize with oligonucleotides 10 attached to nucleic acid detection reagent array 406 after one or more washing steps are performed to wash away non-hybridized nucleic acids from nucleic acid detection reagent array 406 using fluid transfer component 404. As additionally shown, thermal modulator 408 is also operably connected to computer 400. Prior to performing a hybridization assay, target SARS-CoV nucleic acids can be amplified using labeled 15 primer nucleic acids (e.g., primers comprising sequences selected from SEQ ID NOS: 1-12 and 15-24 or complements thereof). The amplicons of these amplification reactions are then typically transferred to nucleic acid detection reagent array 406 using fluid transfer component 404, as described above, to perform the binding assay. In some embodiments, binding assays are performed concurrently with SARS-CoV 20 nucleic acid amplification in thermal modulator 408 using, e.g., molecular beacons, 5'-nuclease probes, or the like that comprise sequences selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. In these embodiments, 25 detector 402 detects detectable signals produced as the amplification reactions are performed using thermal modulator 408.

[0179] Figure 5 schematically shows information appliance or digital device 500 that may be understood as a logical apparatus that can read instructions from media 502 and/or network port 504, which can optionally be connected to server 506 having 30 fixed media 508. Digital device 500 can thereafter use those instructions to direct server or client logic, as understood in the art, to embody aspects of the invention. One type of logical apparatus that may embody the invention is a computer system as

illustrated in 500, containing CPU 510, optional input devices 512 and 514, disk drives 516 and optional monitor 518. Fixed media 502, or fixed media 508 over port 504, may be used to program such a system and may represent a disk-type optical or magnetic media, magnetic tape, solid state dynamic or static memory, or the like. In 5 specific embodiments, the invention may be embodied in whole or in part as software recorded on this fixed media. Communication port 504 may also be used to initially receive instructions that are used to program such a system and may represent any type of communication connection. Optionally, the invention is embodied in whole or in part within the circuitry of an application specific integrated circuit (ACIS) or a 10 programmable logic device (PLD). In such a case, the invention may be embodied in a computer understandable descriptor language, which may be used to create an ASIC, or PLD.

[0180] Figure 5 also includes automatic synthesizer 350, which is operably connected to digital device 500 via server 506. Optionally, automatic synthesizer 520 15 is directly connected to digital device 500. During operation, automatic synthesizer 520 typically receives instructions to synthesize one or more primers or probes that comprise a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID. 20 NOS: 1-12 and 15-24 and the variant, which are included in a data set comprised by, e.g., digital device 500 and/or a computer readable medium, such as fixed media 502 and/or 508.

## XI. KITS

[0181] The nucleic acid detection reagents (e.g., primer nucleic acids, probe 25 nucleic acids, etc.) employed in the methods of the present invention are optionally packaged into kits. As described herein, the nucleic acid detection reagents of the invention detectably bind to target SARS-CoV nucleic acids. In addition, the kits may also include suitably packaged reagents and materials needed for nucleic acid immobilization, hybridization, and/or detection, such solid supports, buffers, enzymes, 30 and nucleic acid standards, as well as instructions for conducting the assay. Optionally, the nucleic acid detection reagents (e.g., oligonucleotides, sequence specific antibodies,

etc.) of the invention are provided already attached or otherwise immobilized on solid supports. As another option, nucleic acid detection reagents are provided free in solution in containers, e.g., for performing the detection methods of the invention in the solution phase. In some of these embodiments, nucleic acid detection reagents of the 5 kits comprise labels and/or quencher moieties, such as when molecular beacons, 5'-nuclease probes, or the like comprise sequences selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant. In certain 10 embodiments, kits further include labeled primers for amplifying target SARS-CoV sequences in a sample.

[0182] The kit also includes one or more of: a set of instructions for contacting the nucleic acid detection reagents with nucleic acids from a sample or amplicons thereof and detecting binding between the nucleic acid detection reagents and SARS-CoV nucleic acids, or at least one container for packaging the nucleic acid detection reagents and the set of instructions. Exemplary solid supports include in the kits of the invention are optionally selected from, e.g., a plate, a microwell plate, a bead, a microbead, a tube (e.g., a microtube, etc.), a fiber, a whisker, a comb, a hybridization chip, a membrane, a single crystal, a ceramic layer, a self-assembling monolayer, or the like. 20

[0183] In some embodiments, the kit further includes at least one primer nucleic acid that is at least partially complementary to at least one segment of a SARS-CoV nucleic acid, e.g., for amplifying a segment of the SARS-CoV nucleic acid. In these embodiments, the kit typically further includes a set of instructions for amplifying one 25 or more subsequences of that nucleic acid with the primer nucleic acids, at least one nucleotide incorporating biocatalyst (e.g., a polymerase), and one or more deoxyribonucleotides. In certain embodiments, the primer nucleic acids comprise at least one label (e.g., a fluorescent dye, a radioisotope, etc.). Suitable labels are described further herein. For example, the primer nucleic acid is optionally conjugated 30 with biotin or a biotin derivative. In these embodiments, the kit typically further includes an enzyme conjugated with avidin or an avidin derivative, or streptavidin or a

streptavidin derivative, e.g., for effecting the detection of binding between the nucleic acid detection reagents of the invention and target nucleic acids. In these embodiments, the kit generally further includes at least one nucleotide incorporating biocatalyst (e.g., a polymerase, a ligase, or the like). In these embodiments, the kit typically also further comprising one or more nucleotides, e.g., for use in amplifying the target nucleic acids. 5 Optionally, at least one of the nucleotides comprises a label. In some of these embodiments, the kits further include at least one pyrophosphatase (e.g., a thermostable pyrophosphatase), e.g., for use in minimizing pyrophosphorolysis, uracil N-glycosylase (UNG) (e.g., a thermostable UNG), e.g., for use in applications where protection 10 against carry-over contamination is desirable.

## XII. EXAMPLES

[0184] It is understood that the examples and embodiments described herein are for illustrative purposes only and are not intended to limit the scope of the claimed invention. It is also understood that various modifications or changes in light the 15 examples and embodiments described herein will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

### EXAMPLE 1: NESTED PCR DETECTION OF SARS CORONAVIRUS

[0185] The SARS Coronavirus Detection RT-PCT Diagnostic is specifically 20 designed for detecting patients suspected of having SARS. Reagents are used for end-point PCR and the products analyzed by agarose gel electrophoresis. Reagents are designed to allow flexibility in the choice of reverse transcriptase and Taq polymerases.

[0186] Reagents provided (10 items):

1. Positive control (2.5 kb DNA fragment comprising the primer sequences) Neat
- 25 2. Positive control 1/10 dilution
3. Positive control 1/100 dilution
4. Negative control A (water or buffer solution alone)
5. Negative control B (cloning vector of the positive control fragment)
6. GISSARSOUT\_F (10  $\mu$ M) (SEQ ID NO: 11)

7. GISSARSOOUT\_R (10 µM) (SEQ ID NO: 12 or 20)
8. GISSARSIN\_F (10 µM) (SEQ ID NO: 13 (= SAR1S; see, Drosten et al. (2003) "Identification of a novel coronavirus in patients with severe acute respiratory syndrome," *N. Engl. J. Med.* 348:1967-1976), SEQ ID NO: 15 or 18)
- 5 9. GISSARSIN\_R (10 µM) (SEQ ID NO: 14 (= SAR1As, see, Drosten et al. (2003), *supra*), SEQ ID NO: 16 or 19)

10. Sterile water for PCR

\* GISSARSOOUT\_R is used both as the 3'-primer for first-strand cDNA synthesis and the first round PCR.

- 10 [0187] Reagents required but not provided:

1. *RNA extraction*

Recommended suppliers: Rneasy® kit (Qiagen)  
or  
TRIZOL® (Invitrogen)

- 15 2. *First-strand cDNA synthesis*

Recommended suppliers: Expand reverse transcriptase (Roche) or  
Superscript II reverse transcriptase (Invitrogen)

- 20 3. *dNTPs Mix*

Recommended suppliers: dNTPs Mix (New England Biolabs)  
or  
dNTPs Mix (Invitrogen)

- 25 4. *Taq Polymerase*

Recommended suppliers: Platinum *Taq Polymerase*  
or  
High Fidelity PCR system (Roche)

- 30 Protocol:

[0188] Samples from patients suspected of having SARS infections, such as blood and tissue need to be handled according to WHO (on the world wide web at [who.int/en/](http://www.who.int/en/) as of April 20, 2004) or CDC (on the world wide web at [cdc.gov/](http://www.cdc.gov/) as of April 20, 2004) recommendations.

1. RNA is extracted by manufacturers' instructions (TRIZOL® or Qiagen kits).
2. First-stand cDNA synthesis is performed on extracted RNA using GISSARSOOUT\_R (2 µl) in a 20 µl reaction volume.
3. First round PCR: The cDNA product (2.5 µl) is used as template using GISSARSOOUT\_F (1.25 µl) and GISSARSOOUT\_R (1.25 µl) in a 25 µl reaction volume.  
5 A 440 bp fragment can be detected at this step if the viral titre of the sample is very high. Note: The control templates (1 µl) can be run along side with every PCR reaction.

#### PCR conditions:

10

94°C for 2 min

15

94°C for 10 sec  
50°C for 30 sec  
72°C for 1 min

72°C for 7 min

4: Second round PCR: Products from the first round PCR (2.5  $\mu$ l) is used as template for the second PCR using GISSARSIN\_F (1.25  $\mu$ l) and GISSARSIN\_R (1.25  $\mu$ l) in a 25  $\mu$ l reaction volume with the same PCR conditions as the first round PCR. Note: The control templates (1  $\mu$ l) can be run along side with every PCR reaction.

5. Products are analyzed on a 1.2 – 1.5% agarose gel. A 120 bp fragment can be detected for positive SARS confirmation. See Figure 6.

25 **EXAMPLE 2: DETECTION OF SARS CORONAVIRUS BY REAL TIME PCR [0189]** SARS-CoV can be detected in a clinical sample using real time PCR rather than the standard RT-PCR method described in Example 1.

The following are mixed in a 20  $\mu\text{L}$  reaction:

10  $\mu$ L 2X Universal PCR Mix (Applied Biosystems)

30 1.8  $\mu$ L 10  $\mu$ M Primer 1 (e.g. SEQ ID NO: 11)

1.8  $\mu$ L 10  $\mu$ M Primer 2 (e.g. SEQ ID NO: 20)

0.5  $\mu$ L, 10  $\mu$ M 5'-nuclease probe (e.g. SEQ ID NO: 21)

4.9  $\mu$ L Water

1.0  $\mu$ L template

[0190] Final concentrations of PCR primers are 900 nM; final concentration of 5'-nuclease probe is 250 nM.

5 [0191] Reactions are preferably run in duplicate.

[0192] PCR cycling is 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C, 15 seconds and annealing and extension at 60°C, 1 minute.

[0193] Fluorescence is monitored as described by the system manufacturer during annealing and extension cycles.

10 [0194] While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques and apparatus described above can be used in various combinations. All publications,

15 patents, patent applications, and/or other documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, and/or other document were individually indicated to be incorporated by reference for all purposes.

JC20 Rec'd PCT/PTO 07 OCT 2005

## APPENDIX

INTERNATIONAL PATENT APPLICATION For

**REAGENTS AND METHODS FOR  
DETECTING SEVERE ACUTE  
RESPIRATORY SYNDROME CORONAVIRUS**

## NOTES

## Detection of Severe Acute Respiratory Syndrome Coronavirus in Blood of Infected Patients

Lisa F. P. Ng,<sup>1</sup> Michelle Wong,<sup>2</sup> Susie Koh,<sup>2</sup> Eng-Eong Ooi,<sup>3</sup> King-Fai Tang,<sup>3</sup> Hoe-Nam Leong,<sup>4</sup> Ai-Ee Ling,<sup>5</sup> Lora V. Agathe,<sup>1</sup> Jenny Tan,<sup>1</sup> Edison T. Liu,<sup>1</sup> Ee-Chee Ren,<sup>1,6</sup> Lee-Ching Ng,<sup>2</sup> and Martin L. Hibberd<sup>\*1</sup>

*Genome Institute of Singapore,<sup>1</sup> DSO National Laboratories,<sup>2</sup> Environmental Health Institute, National Environment Agency,<sup>3</sup> Tan Tock Seng Hospital,<sup>4</sup> Department of Pathology, Singapore General Hospital,<sup>5</sup> and Department of Microbiology, National University of Singapore,<sup>6</sup> Singapore, Republic of Singapore*

Received 3 July 2003/Returned for modification 9 September 2003/Accepted 25 September 2003

Severe acute respiratory syndrome (SARS) has caused major outbreaks worldwide, resulting in an urgent need to obtain sensitive and accurate diagnosis of this disease. PCR-based detection methods were developed for use on a variety of samples, including blood. Eighteen subjects were investigated, and results indicated that blood samples contain sufficient virus for detection by using quantitative real-time PCR.

Severe acute respiratory syndrome (SARS) was first identified in late November 2002 in Guangdong Province, China. In the ensuing months, major outbreaks were reported from Vietnam, Hong Kong, Canada, Singapore, other parts of China, Taiwan, and elsewhere in the world. The disease is unusual in its high level of infectivity, as demonstrated among the health care workers and family members that were in close contact with infected individuals. In addition, it was reported that infected patients do not respond to empirical antimicrobial treatment for acute community-acquired typical or atypical pneumonia (4, 7). The cause of SARS was identified as a novel coronavirus (CoV) (3, 4) because clinical specimens from patients infected with SARS revealed the presence of crown-shaped CoV particles. This new CoV was thus referred to as SARS CoV. The full-length genome sequence of the SARS CoV was reported from different isolates (5, 10), and the genome organization of SARS CoV was found to be similar to that of other CoVs (5, 10).

CoVs are a family of positive-strand RNA-enveloped viruses called *Coronaviridae*, which are now categorized under the newly established order *Nidovirales*. This order comprises the families *Coronaviridae* and *Arteriviridae*. The name *Nidovirales* comes from the Latin word *nidus*, for nest, referring to the 3'-coterminal "nested" set of subgenomic mRNAs produced during viral infection (2). The SARS CoV genome is very large, 29.7 kb (5, 10), and encodes 23 putative proteins. Major structural proteins include nucleocapsid, spike, membrane, and small envelope. Nonstructural proteins include the papain-like proteinase, 3C-like proteinase, RNA-dependent RNA

polymerase (RdRp), helicase, and many other proteins involved with viral replication and transcription (2, 6). In other CoVs, many of the nonstructural proteins are only slightly conserved in the viral sequence, the exception being RdRp, which is highly conserved in many CoVs. In previous studies, primer pairs have been designed against different regions along the SARS CoV (3, 8, 9) and have managed to detect SARS CoV in a variety of clinical samples (3, 8, 9). In one report, the earliest detection observed was in sputum at day 3 (3), while in a different study, detection was found only at day 5 in nasopharyngeal aspirate (NPA), peaking at day 10 (8). In the same study, no association was found with the NPA viral load and clinical progression (8). For this report, we have investigated the use of blood as a means of detecting viral load in SARS patients, a method which in the future may allow improved estimation of disease progression.

Specific primers were designed against the highly conserved polymerase gene of the SARS CoV genome, as sequence comparison among the 14 SARS isolates demonstrated no variations in the SARS CoV RdRp region (10), making the RdRp an ideal region for designing specific diagnostic PCR primers in order to ensure that they will not quickly become obsolete due to sequence mutation. The performance of these primers was assessed in two assays using in vitro-transcribed RNA and virus-spiked samples. In the first PCR assay, 2 µl of RNA was reverse transcribed with Expand reverse transcriptase (Roche, Mannheim, Germany) using primer 5'-GGCATCATCAGAA AGAACATCATCAT-3', thereby generating 20 µl of cDNA. This was amplified with primers 5'-GGTGGGATTATCCAAA TGTGA-3' and 5'-GGCATCATCAGAAAGAACATCATCA T-3' using 2.5 µl in 25-µl reaction mixtures (94°C for 10 s; 35 cycles of 94°C for 10 s, 50°C for 30 s, and 72°C for 1 min; and finally, 72°C for 7 min) with an Expand High-Fidelity PCR system (Roche). Nested primers (5'-ACTATATGTT

\* Corresponding author. Mailing address: Genome Institute of Singapore, 60 Biopolis St., Genome, #02-01, Singapore 138672, Republic of Singapore. Phone: 65-647-88087. Fax: 65-647-89004. E-mail: hibberdm@gis.a-star.edu.sg.

TABLE 1. Investigation of subjects for presence of SARS CoV\*

Subject no.	Diagnosis	Days from onset	Days from illness to discharge or result	Result of quantitative real-time PCR (copies/ml of blood)	Result of gel-based nested PCR
1	Probable SARS	2	9	2,020	+
2	Probable SARS	4	Died	9,100	+
3	Probable SARS	4	>30	3,120	+
4	Probable SARS	5	13	9,600	+
5	Probable SARS	12	18	Negative	-
6	Probable SARS	13	18	Negative	-
7	Probable SARS	14	>30	Negative	-
8	Probable SARS	14	36	Negative	-
9	Non-SARS	0	0	Negative	-
10	Non-SARS	2	1	Negative	-
11	Non-SARS	3	12	Negative	-
12	Non-SARS	10	9	Negative	-
13	Control			Negative	-
14	Control			Negative	-
15	Control			Negative	-
16	Control			Negative	-
17	Control			Negative	-
18	Control			Negative	-

\* Eighteen subjects were tested by both nested PCR and quantitative real-time PCR (LightCycler). Four out of eight subjects with probable SARS cases tested positive. Specimens from non-SARS cases and controls from normal healthy individuals consistently tested negative.

AAACCAGGTGG-3' and 5'-ATTACATTGGCTGTAAAC AGC-3') were used in a second round of PCR, which used 2.5  $\mu$ l of first-round PCR product as a template in a 25- $\mu$ l reaction mixture. The size of this nested PCR product was 110 bp and was resolved in 1.5 to 2% agarose gels. PCR products were sequenced directly to confirm the identity of the products. Results are shown in Table 1.

Next, we used the LightCycler SARS CoV quantification kit (Roche) in a one-step reverse transcriptase PCR for the real-time quantitative PCR that was also designed for the polymerase region and utilized sequence-specific hybridization probes as the detection format. Five microliters of RNA was reverse transcribed and amplified in a 20- $\mu$ l reaction mixture according to the manufacturer's recommendations.

In order to test the sensitivities of this assay, virus grown in Vero E6 cells was harvested, titrated, and spiked into 200  $\mu$ l of Tris-EDTA buffer, sputum, stool suspension, and blood. RNA was extracted from spiked samples by using the HighPure nucleic acid viral kit (Roche) and the QIAamp viral RNA Mini kit (Qiagen) according to the manufacturers' instructions. Comparable results were achieved. Quantitated single-stranded RNA standards (provided by Roche) showed the sensitivity of the assay to be less than 85 copies per reaction (Fig. 1A). The melting curves (Fig. 1A, inset) also confirmed the specificity of the PCR. The internal control showed that there was no inhibition in the samples that had negative signals (Fig. 1B). The detected viral load in the spike samples varied from  $1 \times 10^3$  to  $5 \times 10^3$  copies per reaction (Fig. 1A). This suggests that the extraction protocol used was approximately 10% efficient, with sputum having the greatest efficiency for the clinical samples.

We noted that during the outbreak, most clinical specimens collected for SARS detection (including sputum, NPA, and endotracheal tube samples) involved considerable risk to the health care worker; thus, in order to reduce this risk, we have evaluated the suitability of blood for SARS CoV de-

tection. We investigated 18 subjects comprising eight probable SARS patients and four patients who initially had symptoms similar to SARS but were later diagnosed otherwise. Two of the four subjects who were initially "suspect" for SARS were later diagnosed as having pulmonary tuberculosis and *Escherichia coli* urinary tract infection. The diagnosis of the other two patients was not determined. Four out of eight probable SARS patients had detectable virus in their blood, while the four non-SARS patients and six control samples from healthy individuals showed no virus (Fig. 2 and Table 1). In addition, control samples spiked with other human CoVs showed no signals (data not shown). Once again, melting-curve analysis confirmed the specificity of the PCR products (Fig. 2, inset). Viral load in the four positive SARS patients varied from  $2 \times 10^3$  copies to  $1 \times 10^4$  copies per ml of blood (Fig. 2; Table 1). It is unclear from these data whether this variation predicts clinical outcome; however, work from other studies has shown that blood viral load is a good indicator of disease progression (1). The two PCR methods described herein showed equal sensitivity in terms of detection. However, the gel-based assay method is laborious and nonquantitative.

We detected virus in blood at 2 days after the onset of symptoms, which is earlier than previously reported (3, 8, 9). Although it is important to further define the window period for detection of SARS CoV in blood, it is interesting to speculate that the time course of SARS CoV viremia may be relatively short, with our data suggesting that, at days 12 to 14, viral load is too low to detect (Table 1). More patients and sequential sampling would be required to confirm this supposition.

In conclusion, we have shown that the SARS CoV viral load can be determined in patients' blood by using PCR methods, and our data suggest that SARS CoV can be detected early in blood, i.e., within the first week of symptom onset. This fact allows early diagnosis and determination of viral load, both of

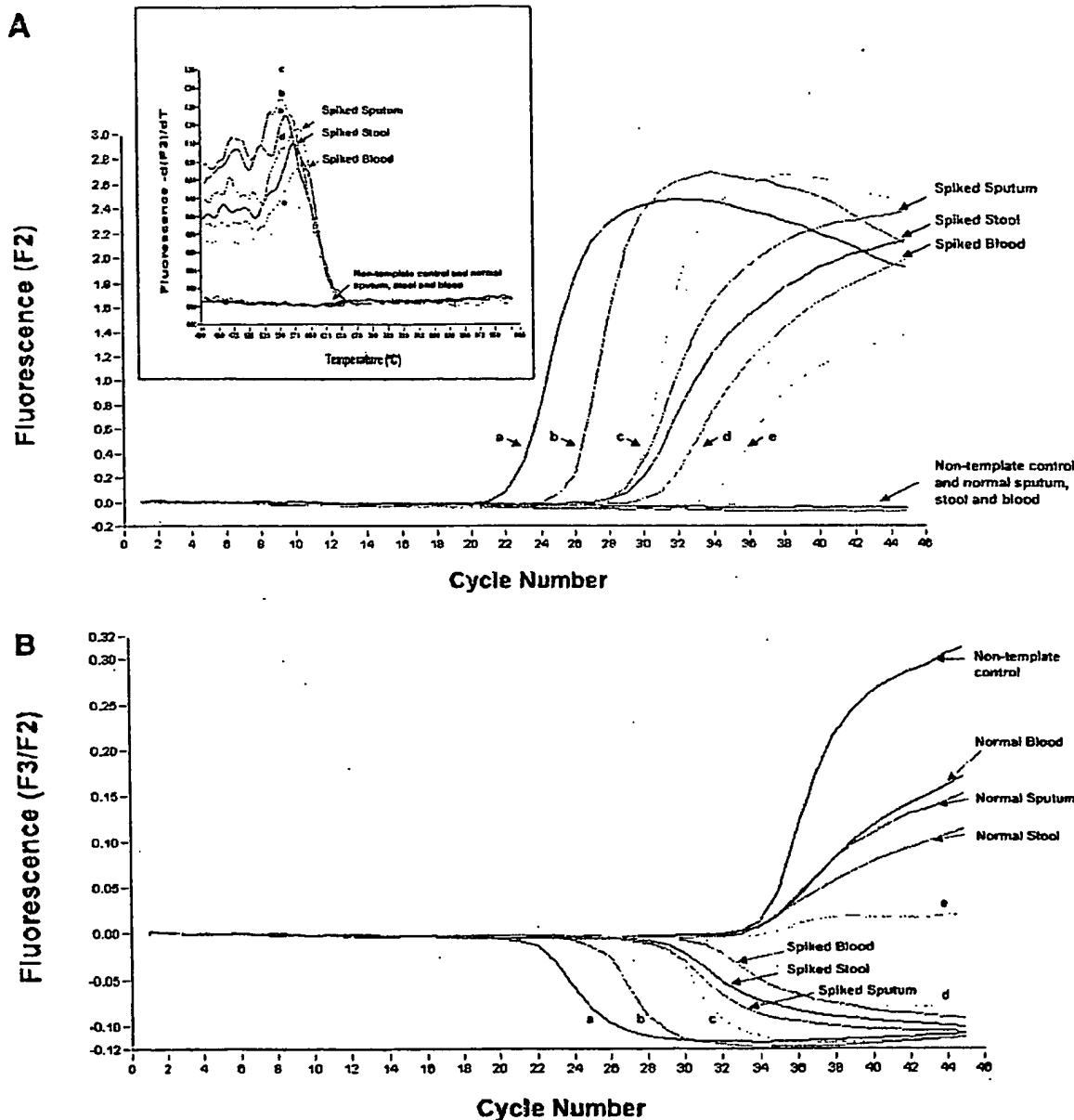


FIG. 1. Detection of SARS CoV by real-time quantitative PCR. (A) Amplification of single-stranded RNA standards (indicated as a to e) and RNA extracted from sputum, stool, and blood spiked with virus grown in Vero E6 cells. The x axis denotes the cycle number of the quantitative PCR assay, and the y axis denotes fluorescence intensity (F2) over the background level. The nontemplate control (water) is indicated. The viral load is indicated as the number of copies per reaction: spiked sputum,  $5 \times 10^6$ ; spiked stool,  $4 \times 10^5$ ; spiked blood,  $1 \times 10^3$ . RNA standards were as follows: (a)  $1 \times 10^6$  copies per reaction, (b)  $9.5 \times 10^5$  copies per reaction, (c)  $8.7 \times 10^3$  copies per reaction, (d)  $1.1 \times 10^3$  copies per reaction, and (e)  $8.5 \times 10^2$  copies per reaction. The inset graph represents melting-curve analysis of the PCR products. Signals from RNA standards (a to e), spiked samples, normal samples, and nontemplate control (water) are shown. The x axis denotes the temperature (°C), and the y axis denotes the fluorescence intensity (F2) over the background level. (B) Detection of the internal control in fluorometer channel F3 in parallel with the simultaneous amplification of the RNA standards (a to e), spiked samples, and nontemplate control is shown.

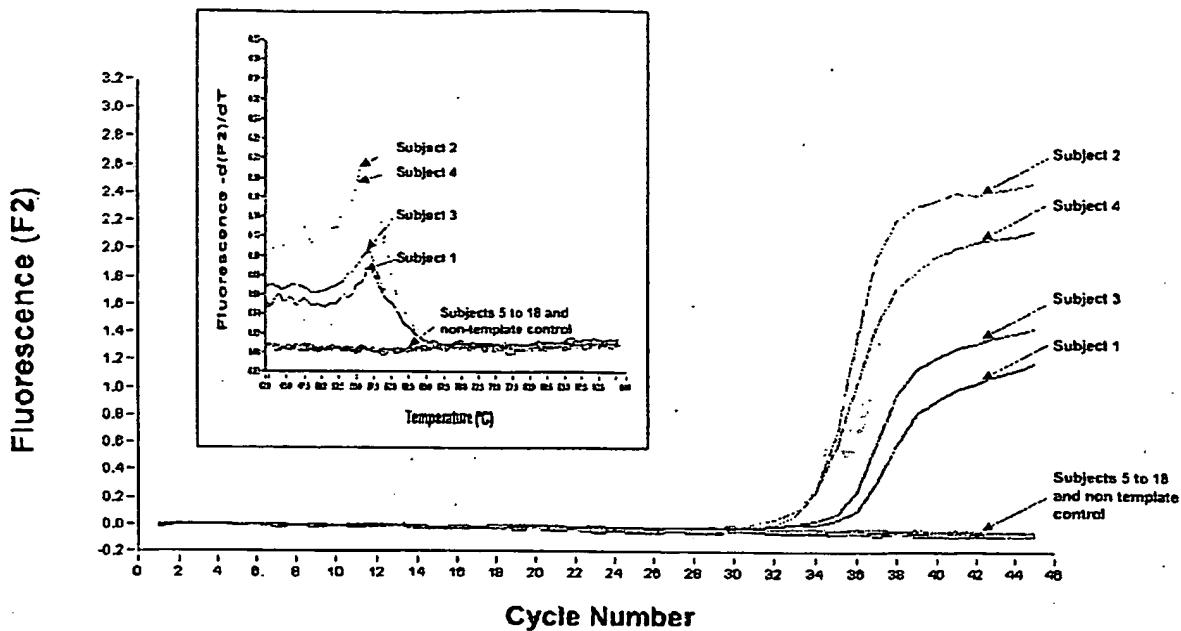


FIG. 2. Detection of SARS CoV from 18 subjects by real-time quantitative PCR. Data for RNA from 18 subjects are shown. The viral load is indicated as number of copies per ml of blood. Subject 1, 2,020 copies; subject 2, 9,100 copies; subject 3, 3,120 copies; subject 4, 9,600 copies. A sample would be considered positive if it generated a typical amplification curve within the 45 cycles. Negative signals from subjects 5 to 18 and the nontemplate control (water) are shown. The inset graph shows melting-curve analysis of the PCR products from the 18 subjects. The x axis denotes the temperature (°C), and the y axis denotes the fluorescence intensity (F2) over the background level.

which are useful in clinical and public health management settings. Further work is under way to further define the role of blood in diagnosis.

This work was funded by the Agency of Science and Technology (A\*STAR), Singapore, Republic of Singapore.

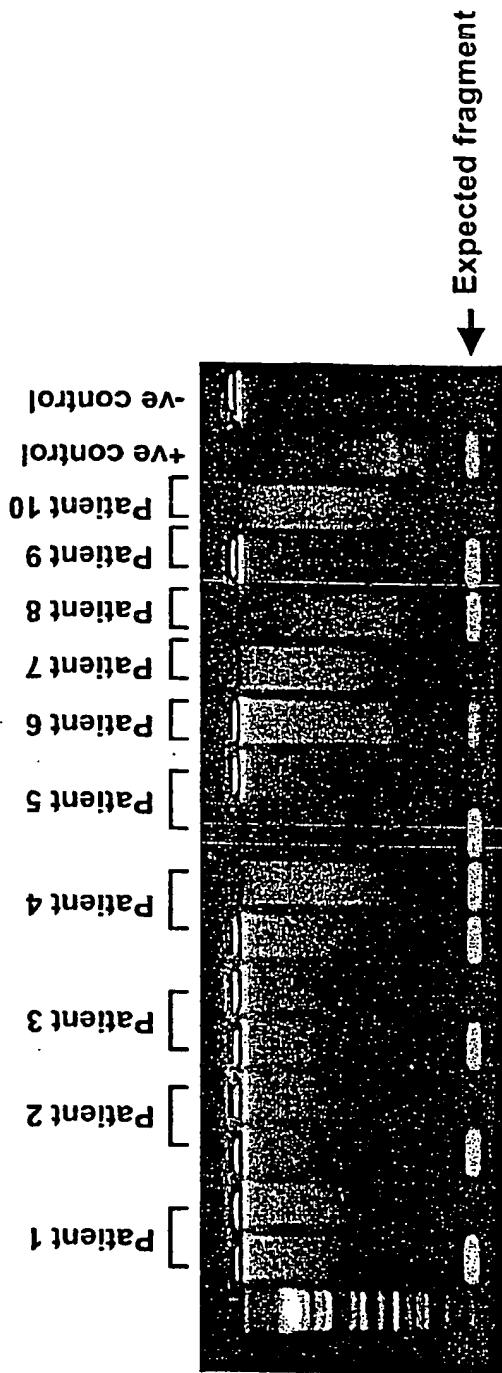
We thank Loh Gek Kee and Tan Yian Kim (DSO National Laboratories), Patricia Tay and Khoo Chen Ai (Genome Institute of Singapore), Kamal Singh (National University Hospital), and the doctors and staff at Tan Tock Seng Hospital.

#### REFERENCES

1. Condotti, D., J. Temple, F. Sarkodie, and J. P. Allain. 2003. Frequent recovery and broad genotype 2 diversity characterize hepatitis C virus infection in Ghana, West Africa. *J. Virol.* 77:7914-7923.
2. Cowanagh, D. 2003. Nidovirales: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch. Virol.* 14:629-633.
3. Drosten, C., S. Günther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Foucier, A. Berger, A. M. Burguiere, J. Cinni, M. Eickmann, N. Escrivio, K. Grywna, S. Kramme, J. C. Manuguerra, S. Müller, V. Rücker, M. Stürmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H. Schmitz, and H. W. Doerr. 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* 348:1967-1976.
4. Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rosa, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, L. J. Anderson, and the SARS Working Group. 2003. A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* 348:1953-1966.
5. Morra, M. A., S. J. M. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Andonov, J. I. Artsob, N. Bastien, K. Bernard, T. F. Booth, D. Bourns, M. Czub, M. Drebout, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Myers, A. Kabani, Y. Li, S. Normand, U. Ströh, G. A. Tipples, S. Tyler, R. Vugrin, D. Ward, B. Watson, R. C. Brumham, M. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper. 2003. The genome sequence of the SARS-associated coronavirus. *Science* 300:1399-1404. (First published online 1 May 2003; 10.1126/science.1085963.)
6. Ng, L. P. P., and D. X. Liu. 2002. Membrane association and dimerization of a cysteine-rich, 16-kilodalton polypeptide released from the C-terminal region of the coronavirus infectious bronchitis virus 1a polyprotein. *J. Virol.* 76:6257-6267.
7. Peiris, J. S. M., S. T. Lai, L. L. Poon, Y. Guan, L. Y. Yam, W. Lim, J. Nicholls, W. K. Yee, W. W. Yau, M. T. Cheung, V. C. Cheng, K. H. Chan, D. N. Tsang, R. W. Yung, T. K. Ng, K. Y. Yuen, and the SARS Study Group. 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361:1319-1325.
8. Peiris, J. S. M., C. M. Chu, V. C. C. Cheng, K. S. Chan, J. F. N. Hung, L. L. Poon, K. I. Law, B. S. F. Tong, T. Y. W. Hon, C. S. Chan, K. H. Chan, J. S. C. Ng, B. J. Zheng, W. L. Ng, R. W. M. Lai, Y. Guan, K. Y. Yuen, and the HKU/UCH SARS Study Group. 2003. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet* 361:628-633.
9. Poon, L. L. M., O. K. Wong, W. Luk, Y. Y. Kwok, J. S. M. Peiris, and Y. Guan. 2003. Rapid diagnosis of a coronavirus associated with severe acute respiratory syndrome (SARS). *Clin. Chem.* 49:953-955.
10. Guan, Y. J., C. L. Wei, A. E. Ling, V. B. Vega, H. Thoresen, S. Y. Se Thoe, J. M. Chia, P. Ng, K. P. Chiu, L. Lim, T. Zhan, K. P. Chan, L. L. E. Oon, M. L. Ng, Y. S. Leo, L. F. P. Ng, E. C. Ren, L. W. Stanton, P. M. Long, and E. T. Liu. 2003. Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. *Lancet* 361:1779-1785.

**Nested PCR on SARS suspect patients**

1<sup>st</sup> pcr: GISSARSOUT-F/R  
2<sup>nd</sup> pcr: GISSARSIN-F/R



Annex: Protocol for LightCycler RNA Master SYBR GreenProtocol:

1. 2  $\mu$ l of sample RNA is used for each reaction. Note: The positive control template (1  $\mu$ l) should be run along. A negative control should also be run with the samples.
2. Reagent preparation should be done in either 0.2 ml or 1.5 ml reaction tubes on ice added in the following order:

Component	Volume
Sterile water	X $\mu$ l
Mn(OAc) <sub>2</sub> stock solution, 50 mM	1.3 $\mu$ l
LC F / R Mix	2 $\mu$ l
Template	X $\mu$ l
LightCycler RNA Master SYBR Green I	7.5 $\mu$ l

Mix gently and transfer reaction mix to glass capillaries.

**Program 1: Reverse Transcription**

The recommended T°C for the RT step is 61°C.

<u>Cycle Prog Data</u>	<u>Value</u>
Cycles	1
Analysis Mode	None
Temperature Targets	Segment 1
Target T°C	61
Incubation time	20 min
T°C transition rate	20
Secondary target T°C	0
Step size (°C)	0.0
Step Delay (cycles)	0
Acquisition mode	None

**Program 2: Denaturation**

<u>Cycle Prog Data</u>	<u>Value</u>
Cycles	1
Analysis Mode	None
Temperature Targets	Segment 1
Target T°C	95
Incubation time	10 sec
T°C transition rate	20
Secondary target T°C	0
Step size (°C)	0.0
Step Delay (cycles)	0

**Program 3: Acquisition mode** **Amplification** **None**

<u>Cycle Prog Data</u>		<u>Value</u>
Cycles		40
Analysis Mode		Quantification
Temperature Targets	Segment 1	Segment 2
Target T°C	95	50
Incubation time	16 sec	16sec
T°C transition rate	20	20
Secondary target T°C	0	0
Step size (°C)	0.0	0.0
Step Delay (cycles)	0	0
Acquisition mode	None	Single

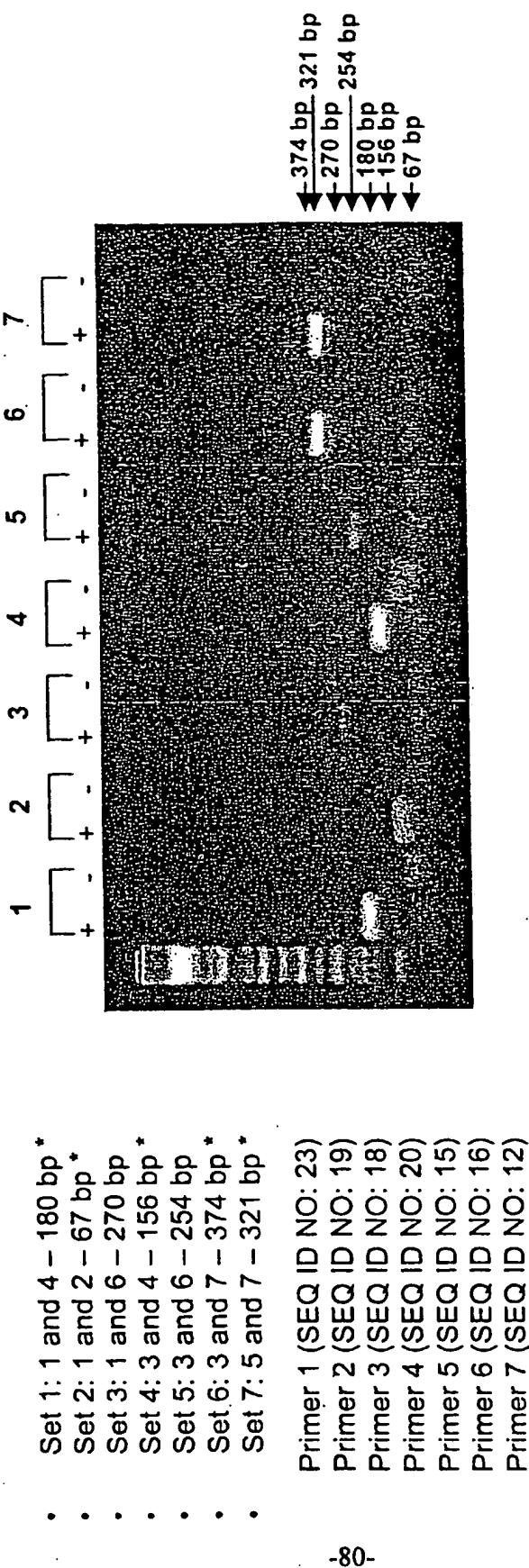
**Program 4: Melting Curve Analysis**

<u>Cycle Prog Data</u>		<u>Value</u>
Cycles		1
Analysis Mode		Melting Curve
Temperature Targets	Segment 1	Segment 2
Target T°C	95	60
Incubation time	5 sec	15 sec
T°C transition rate	20	20
Secondary target T°C	0	0
Step size (°C)	0.0	0.0
Step Delay (cycles)	0	0
Acquisition mode	None	Cont.

**Program 5: Cooling**

<u>Cycle Prog Data</u>		<u>Value</u>
Cycles		1
Analysis Mode		None
Temperature Target		Segment 1
Target T°C		40
Incubation time		30 sec
T°C transition rate		20
Secondary target T°C		0
Step size (°C)		0.0
Step Delay (cycles)		0
Acquisition mode		None

**Annex: LightCycler RNA Master SYBR Green Detection on different primer combinations**



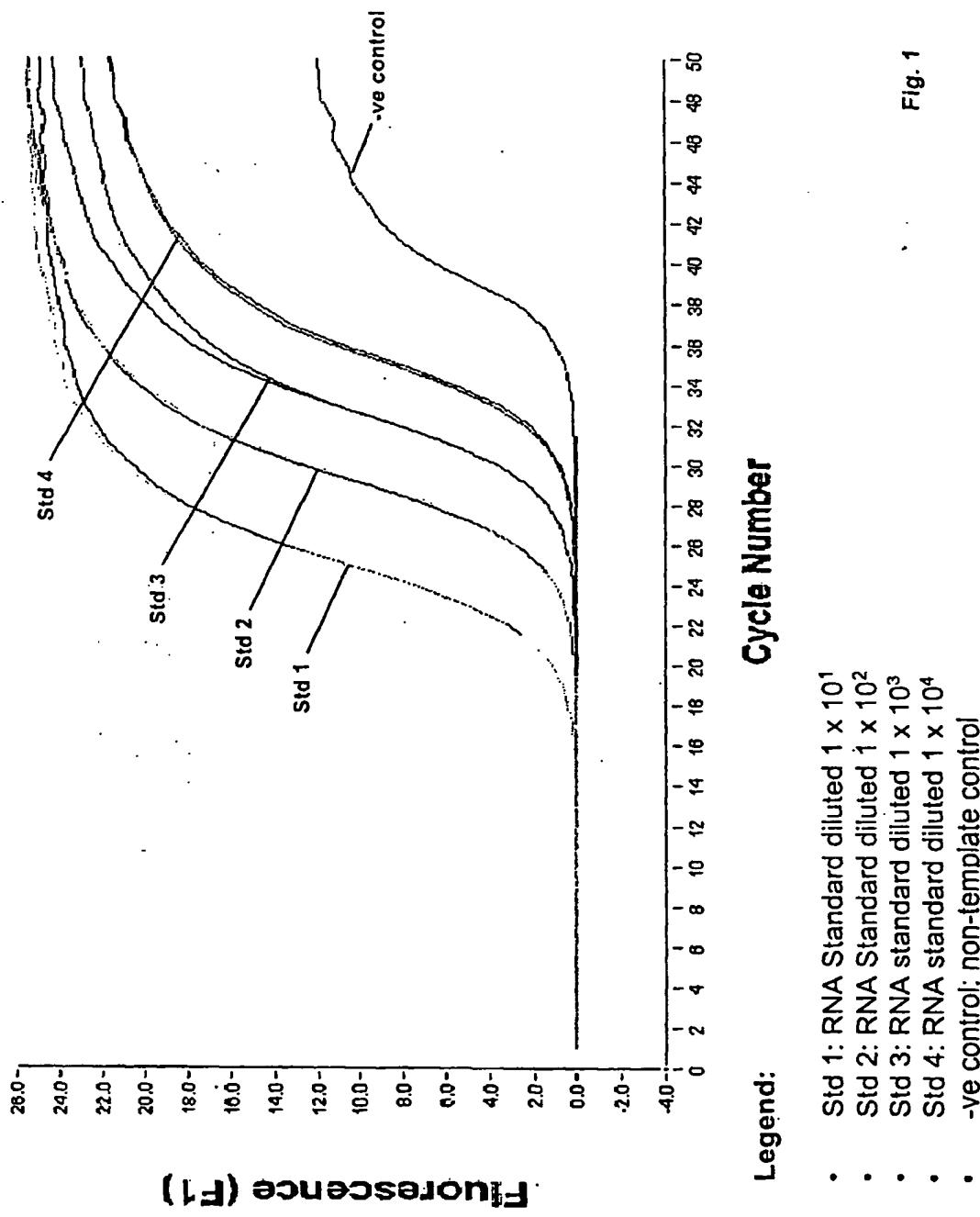
The LightCycler RNA Master SYBR Green 1 (3064760) was used to test the sensitivity of the primers described herein. The template used was RNA isolated from Vero isolates. For primer set 1, 4, 6 and 7 signals showed up between cycles 17 to 20. The same RNA template was also done on the Artus kit and signals showed up only after cycle 23.

Program 1: Reverse Transcription  
Program 2: Denaturation – Incubation time changed to 10 sec  
Program 3: Amplification – Incubation time changed to 16 sec

**SUBSTITUTE SHEET (RULE 26)**

**Annex: Results of LightCycler RNA Master SYBR Green with GIS  
SARS Primers**

**RNA isolated from Serum "spike" with virus obtained Vero isolates**



Melting curve for RNA isolated from "spike" serum

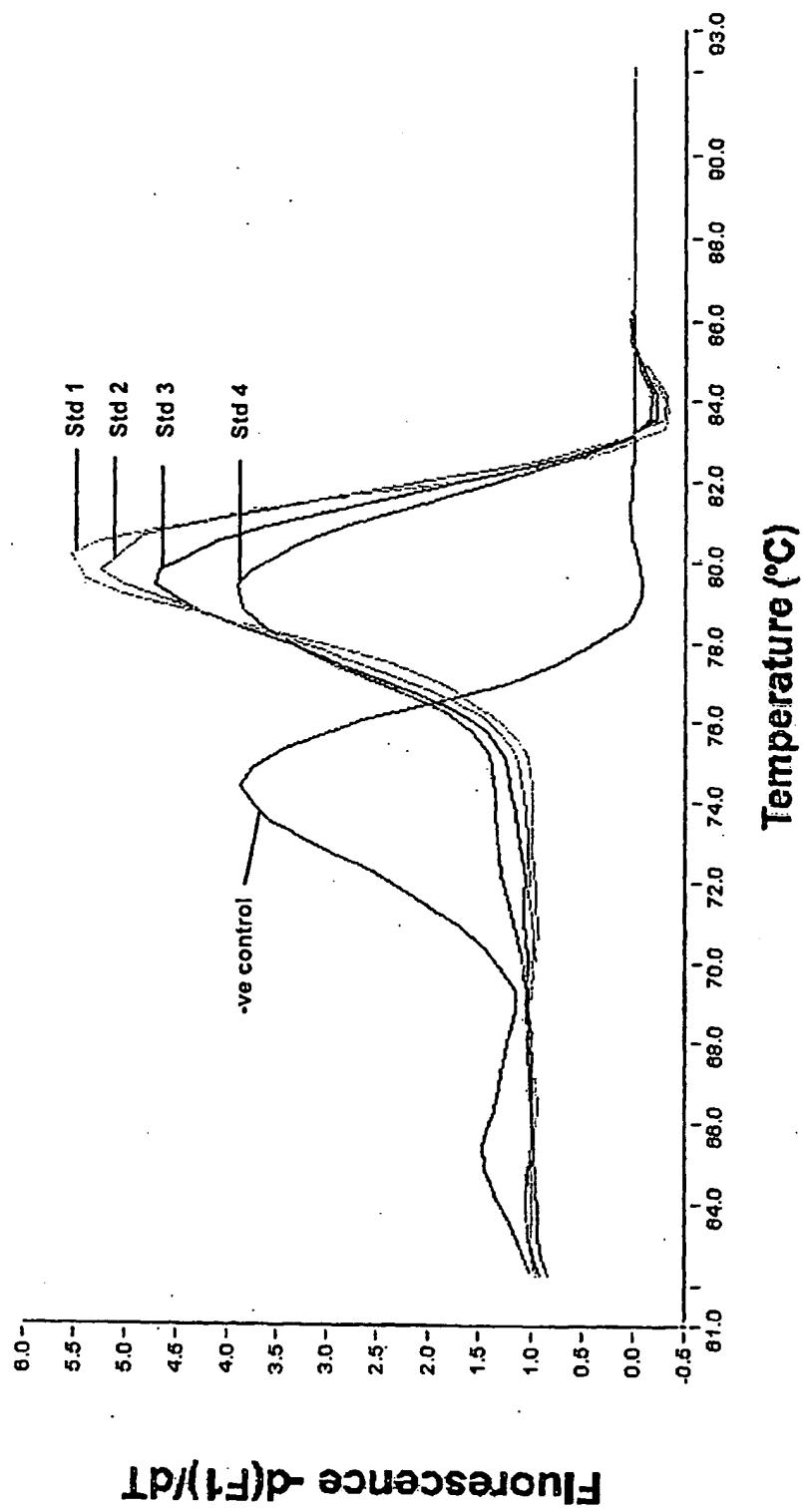


Fig. 1a

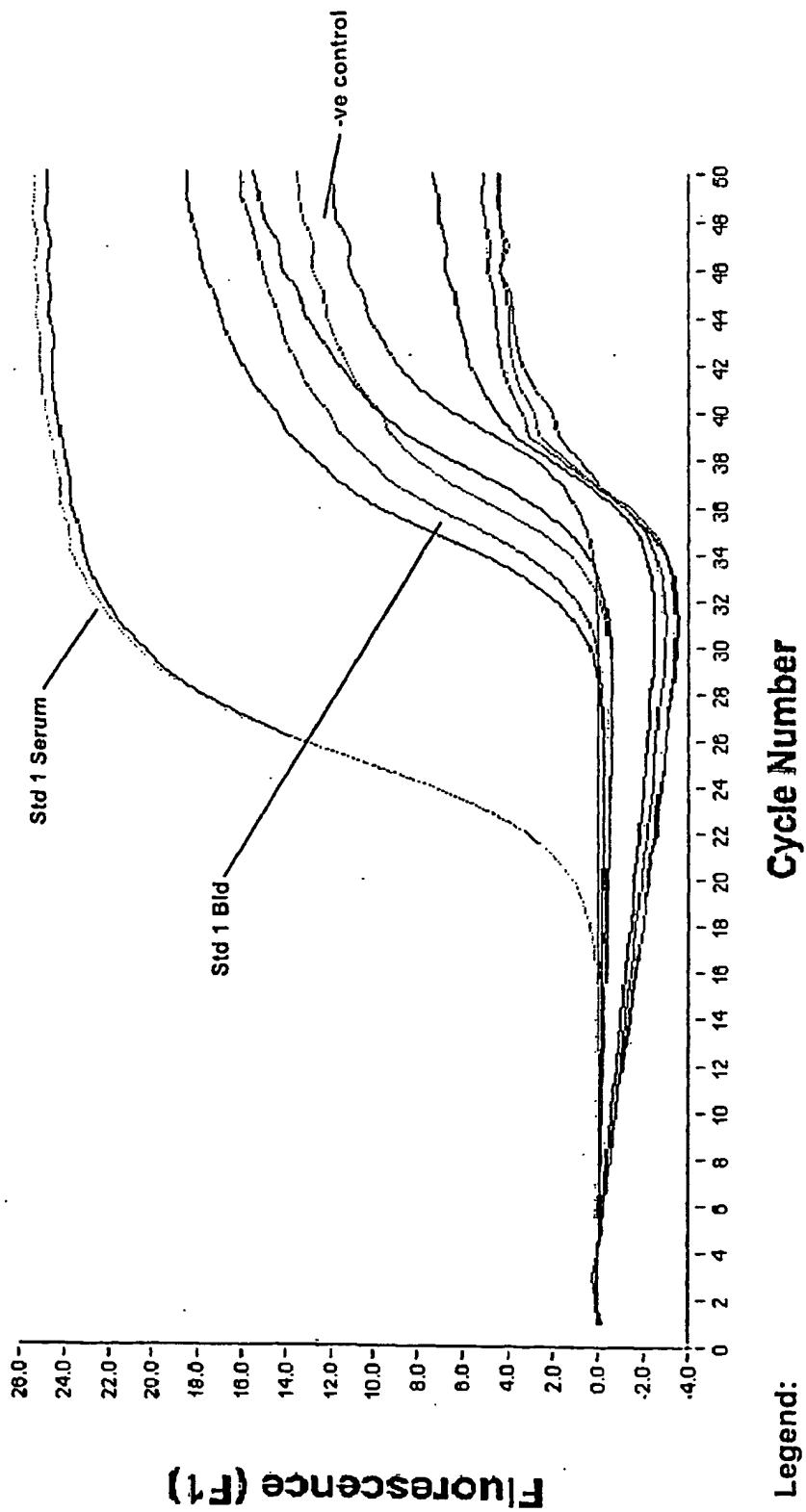
**RNA isolated from whole blood "spike" with virus obtained Vero isolates**

Fig. 2

**SARS-CoV Detection RT-PCR Diagnostic Using 5'-Nuclease Probe on the LightCycler****Specifications:**

The SARS-CoV Detection RT-PCR is specifically designed for detecting patients suspected of having SARS using the LightCycler RNA Master Hybridization Probes ("Hot Start") using a specially designed 5'-Nuclease Probe. This system has been adapted to be used in the LightCycler. Our results to date have been performed on RNA extracted from both whole blood (Trizol) and serum (Qiagen viral kit). Data showed that both extraction methods are fairly comparable, indicating that Trizol do not inhibit the PCR reactions. For sensitivity, this system is able to detect as low as 10 copies at cycle 35.

**Protocol:**

1. Reagent preparation should be done in 1.5 ml reaction tubes on ice added in the following order:

Component	Volume
Sterile water	X $\mu$ l
Mn(OAc) <sub>2</sub> stock solution, 50 mM	1.3 $\mu$ l
LC-ProbeNPrimer Mix	2.6 $\mu$ l
LightCycler RNA Master Hybridization probes	7.5 $\mu$ l

Mix gently and transfer reaction mix to glass capillaries.

2. 5  $\mu$ l of sample RNA is added to each glass capillary.
3. The cycle program is as shown.

**Program 1-Reverse Transcription**

Cycle Program Data	Value
Cycles	1
Analysis Mode	None
Temperature Targets	Segment 1
Target T°C	61
Incubation time	20 min
T°C transition rate (°C/s)	20.0
Secondary Target T°C	0
Step Size (°C)	0.0
Step Delay (cycles)	0
Acquisition Mode	None

**Program 2-Denaturation**

Cycle Program Data	Value
Cycles	1
Analysis Mode	None

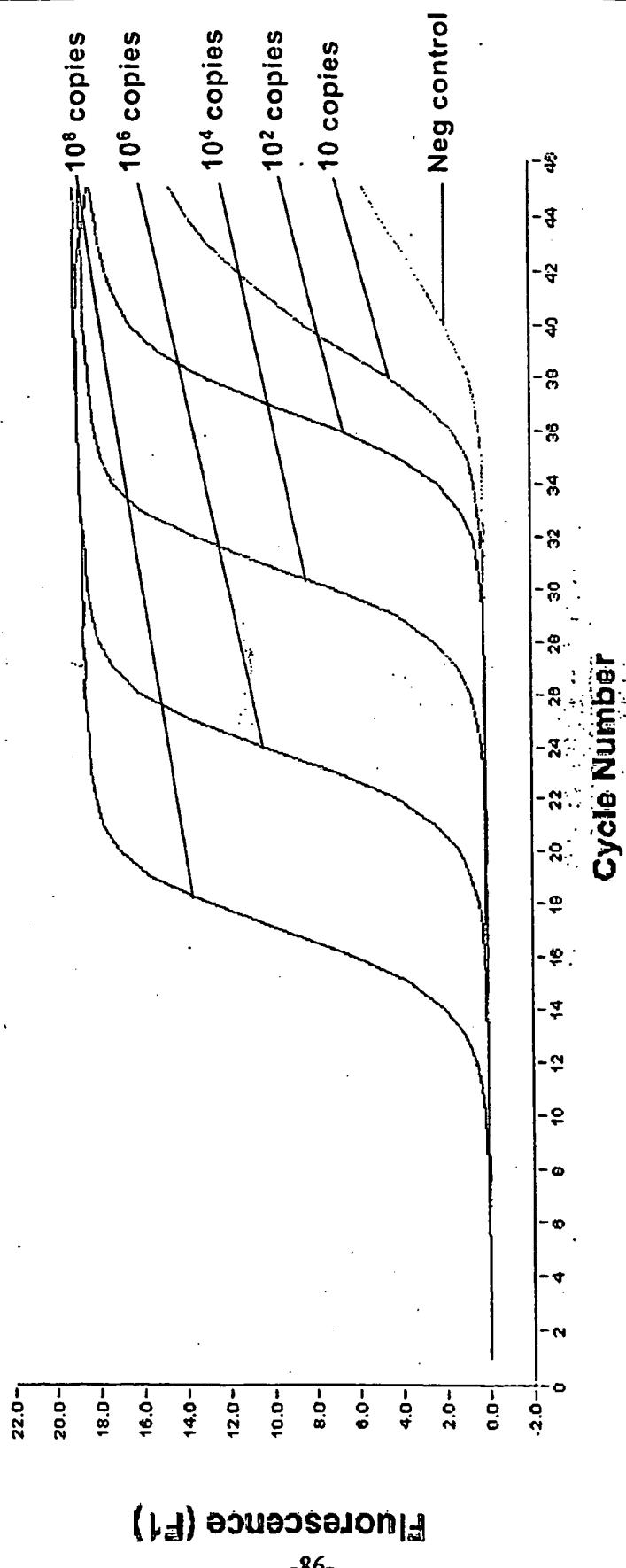
<b>Temperature Targets</b>	<b>Segment 1</b>
Target T°C	95
Incubation time	2 min
T°C transition rate (°C/s)	20.0
Secondary Target T°C	0
Step Size (°C)	0.0
Step Delay (cycles)	0
Acquisition Mode	None

**Program 3-Amplification**

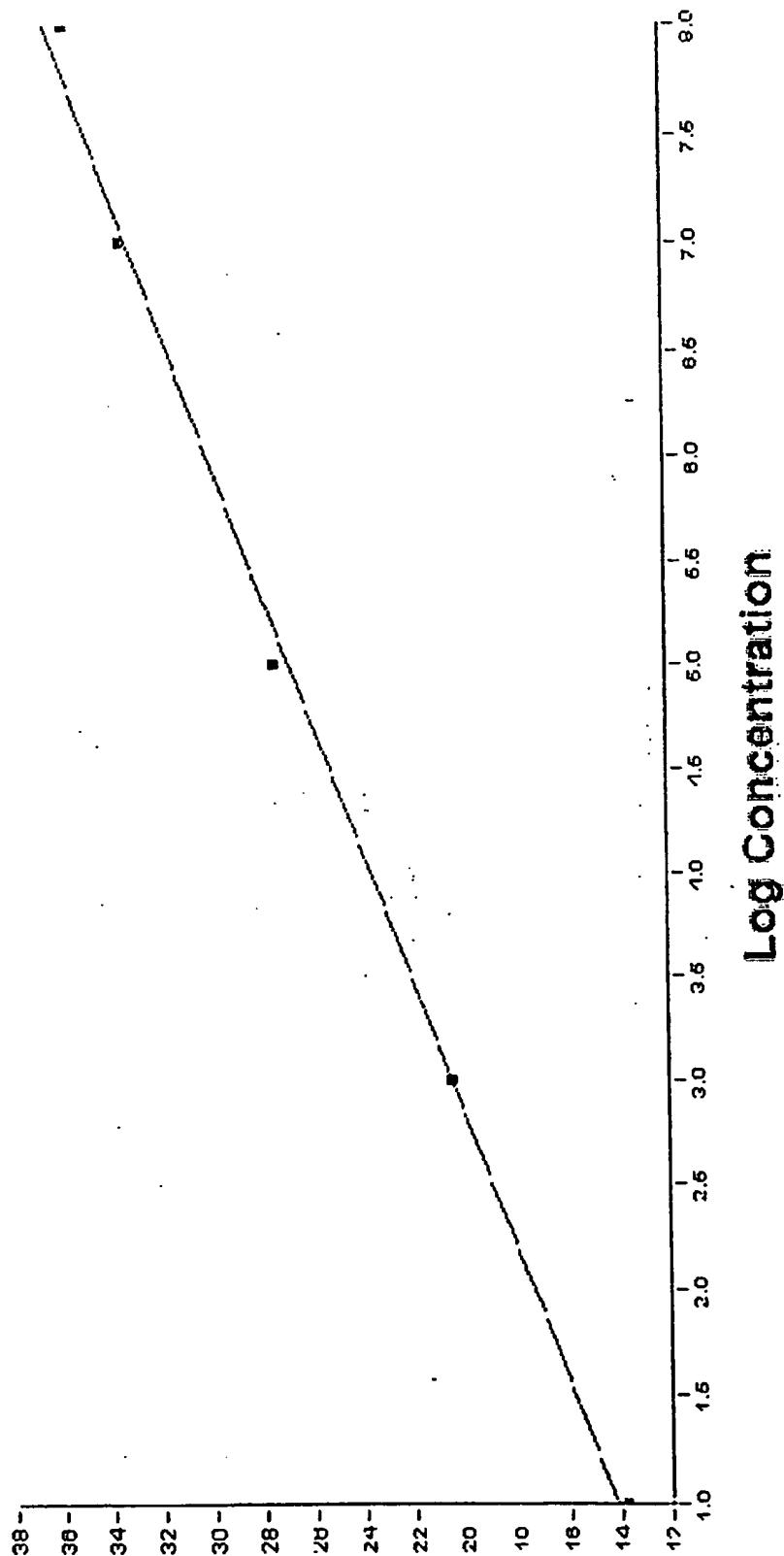
<b>Cycle Program Data</b>	<b>Value</b>
Cycles	1
Analysis Mode	Quantification
<b>Temperature Targets</b>	<b>Segment 1</b> <b>Segment 2</b>
Target T°C	95      60
Incubation time	5 sec      1 min
T°C transition rate (°C/s)	20.0      20.0
Secondary Target T°C	0      0
Step Size (°C)	0.0      0.0
Step Delay (cycles)	0      0
Acquisition Mode	None      Single

**Program 4-Cooling**

<b>Cycle Program Data</b>	<b>Value</b>
Cycles	1
Analysis Mode	None
<b>Temperature Targets</b>	<b>Segment 1</b>
Target T°C	40
Incubation time	30 sec
T°C transition rate (°C/s)	20.0
Secondary Target T°C	0
Step Size (°C)	0.0
Step Delay (cycles)	0
Acquisition Mode	None

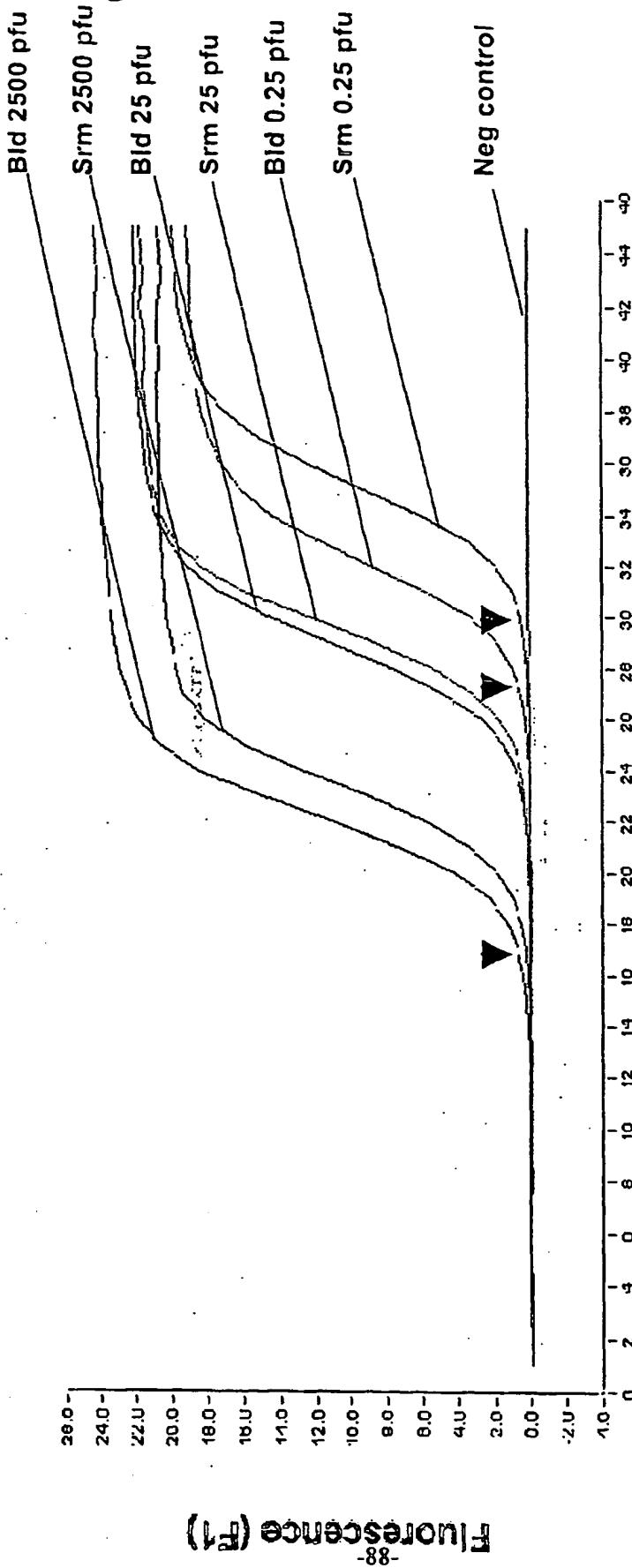
Annex: Results of 5'-Nuclease Probe in LightCycler using GISSARS PrimersGISS primers and hybridisation probe – LightCycler  
(RNA template)

Standard Curve on RNA template



Cycle Number

## GIS primers and hybridisation probe – LightCycler

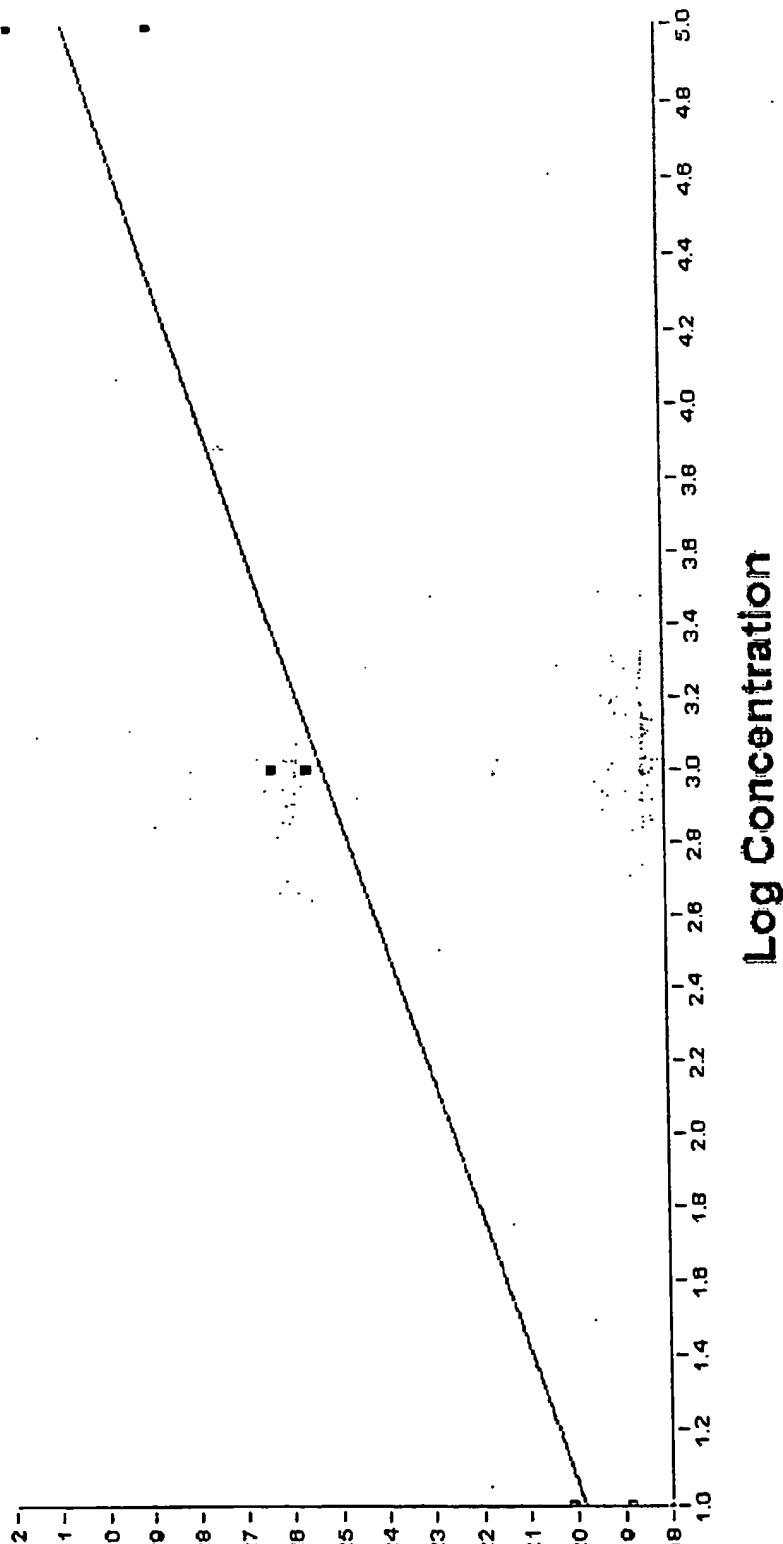


### Abbreviations:

- Bld: RNA isolated from "spike" blood
- Srm: RNA isolated from "spike" serum
- Neg: No template negative control

Signals started showing at cycle 17 and  
can detect 0.25 particles from blood at  
Cycle 27!

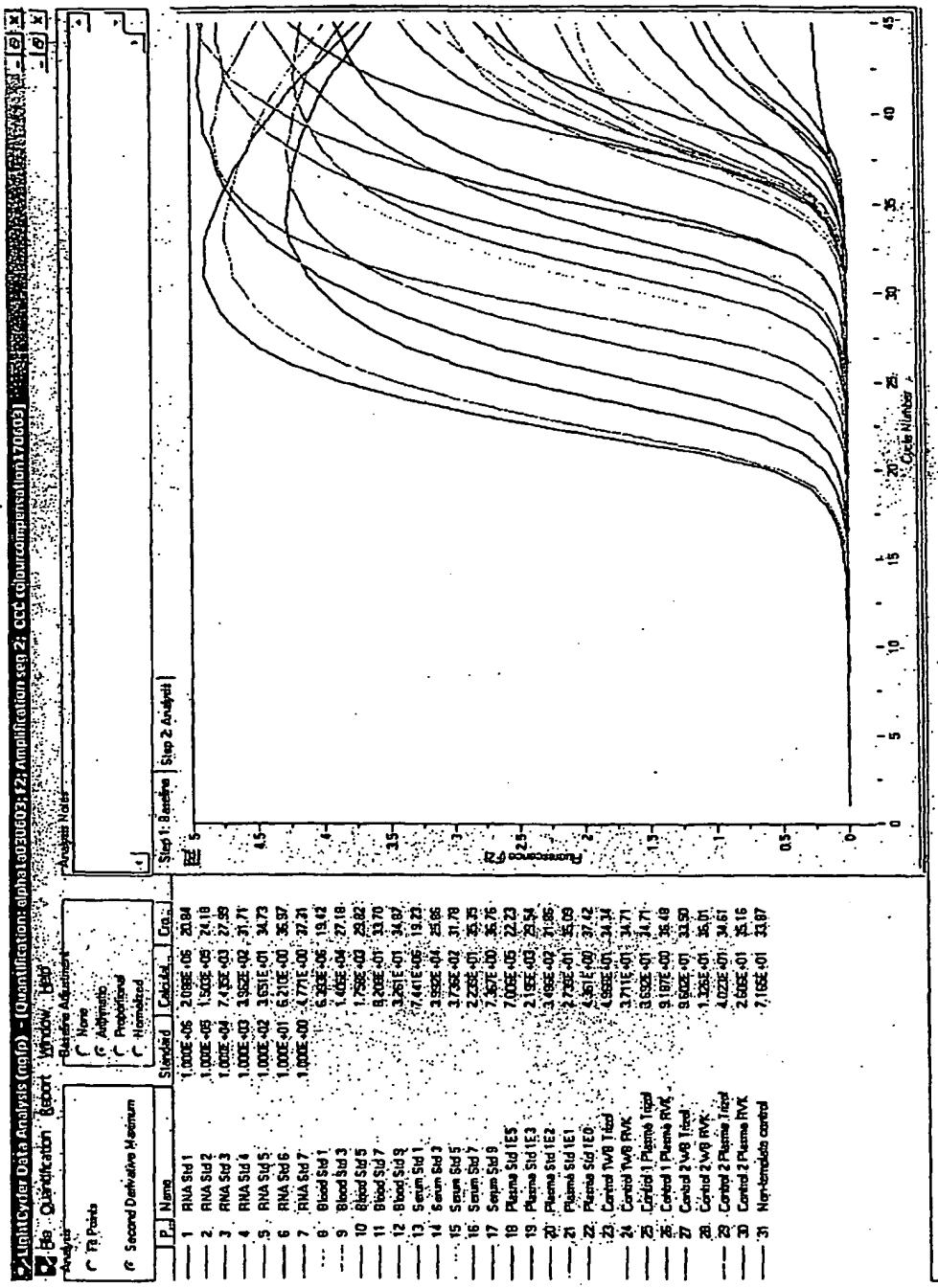
**Standard Curve on Blood and Serum RNA**



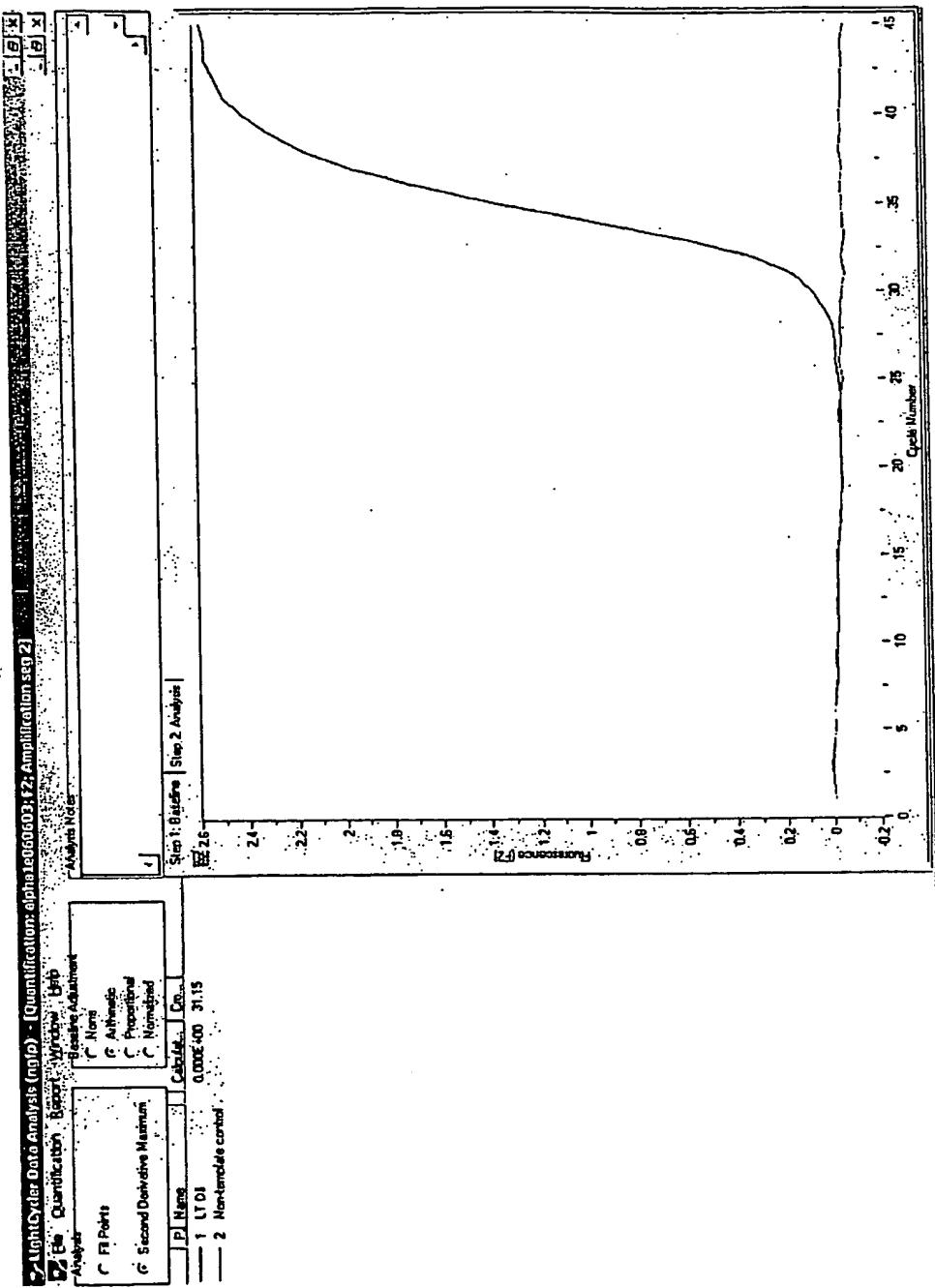
Cycle Number

-89-

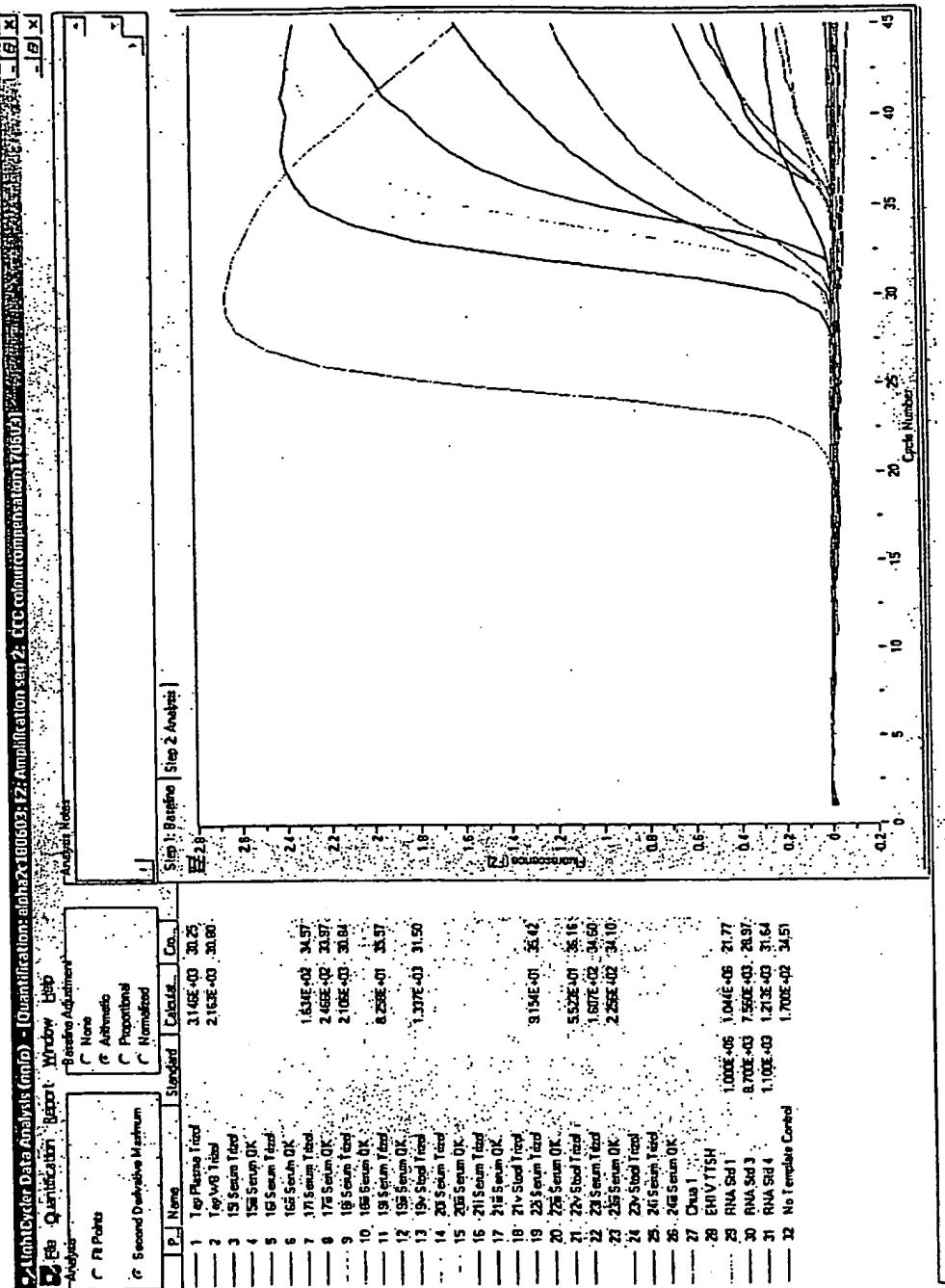
## Test with RNA Standards and Test Samples



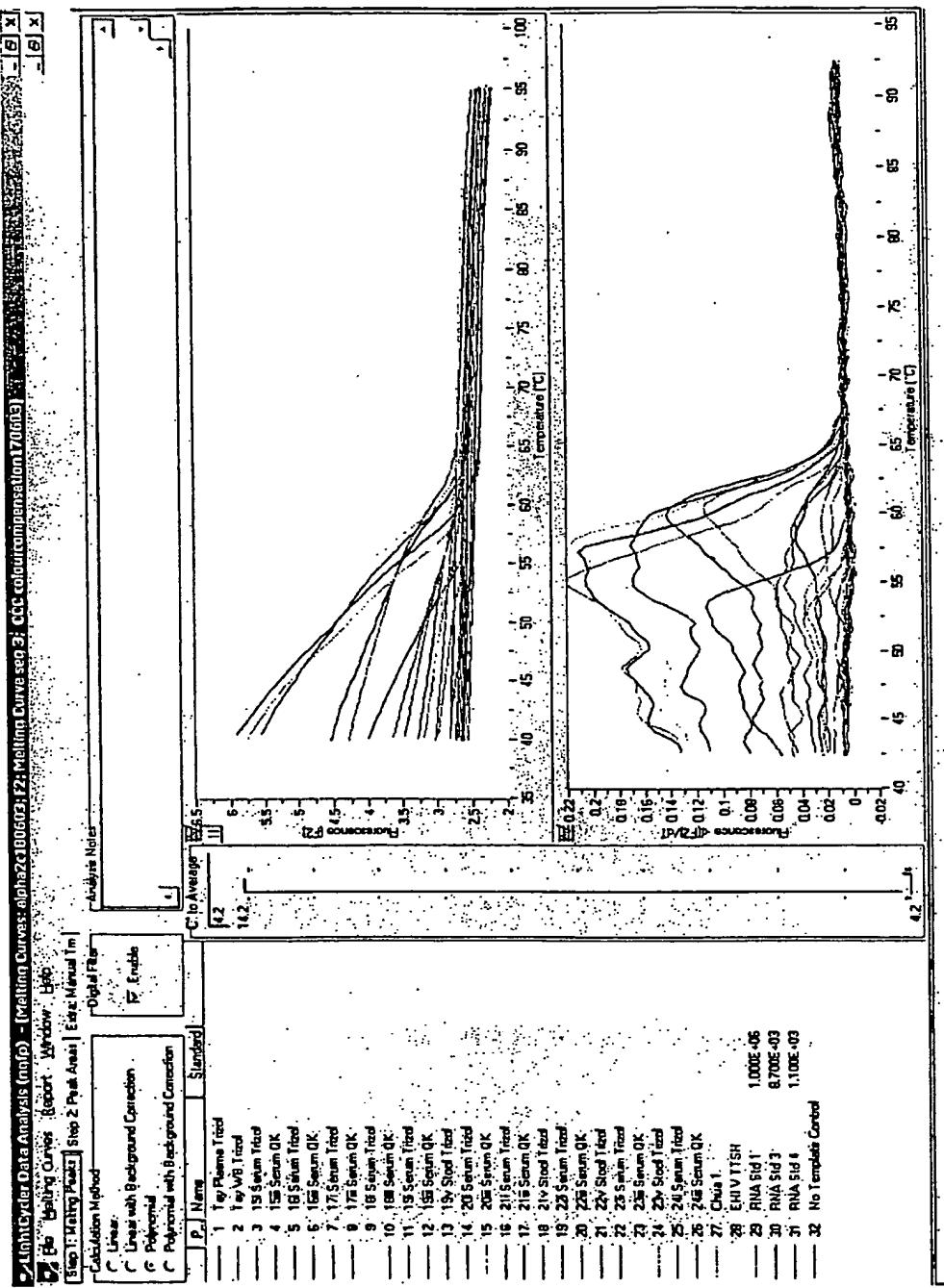
## Test with Positive Control Sample



Test with RNA Standards and Test Samples:  
(Quantitation curves indication viral load)



**Test with RNA Standards and Test Samples:**  
**(Melting curves indicating specificity of PCR)**



## CLAIMS

WHAT IS CLAIMED IS:

1. An oligonucleotide consisting of a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24 and complements thereof.
- 5 2. An oligonucleotide comprising a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24 and complements thereof, which oligonucleotide has 100 or fewer nucleotides.
3. The oligonucleotide of claim 2, wherein the oligonucleotide comprises a probe nucleic acid or a primer nucleic acid.
- 10 4. The oligonucleotide of claim 2, wherein the oligonucleotide comprises at least one modified nucleotide.
5. The oligonucleotide of claim 2, wherein the oligonucleotide comprises at least one label and/or at least one quencher moiety.
6. The oligonucleotide of claim 2, wherein the oligonucleotide has 40 or fewer 15 nucleotides.
7. An oligonucleotide comprising a nucleic acid having at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24 or a complement thereof, which oligonucleotide has 100 or fewer nucleotides.
8. The oligonucleotide of claim 7, wherein the nucleic acid has at least 95% sequence identity to one of SEQ ID NOS: 1-12 and 15-24 or the complement thereof.
- 20 9. The oligonucleotide of claim 7, wherein the oligonucleotide comprises at least one modified nucleotide.
10. The oligonucleotide of claim 7, wherein the oligonucleotide comprises at least one label and/or at least one quencher moiety.
- 25 11. The oligonucleotide of claim 7, wherein the oligonucleotide has 40 or fewer nucleotides.
12. The oligonucleotide of claim 7, wherein the oligonucleotide comprises at least one conservatively modified variation.

13. A method of detecting a severe acute respiratory syndrome coronavirus in a sample, the method comprising:

(a) contacting nucleic acids from the sample with at least one primer nucleic acid comprising at least one nucleic acid selected from the group consisting of:

5. SEQ ID NOS: 1-12 and 15-24 and complements thereof in at least one nucleic acid amplification reaction; and,

(b) detecting the nucleic acids and/or one or more amplicons thereof from the nucleic acid amplification reaction during or after (a), thereby detecting the severe acute respiratory syndrome coronavirus in the sample.

10 14. The method of claim 13, wherein at least one of the primer nucleic acids comprises a modified primer nucleic acid.

15. The method of claim 13, wherein at least one of the amplicons is about 440 nucleotides in length.

16. The method of claim 13, wherein at least one round of the nucleic acid amplification reaction is performed using primer nucleic acids comprising sequences selected from SEQ ID NOS: 11 or 22 and SEQ ID NOS: 12 or 20.

17. The method of claim 13, wherein at least one round of the nucleic acid amplification reaction is performed using primer nucleic acids comprising sequences selected from SEQ ID NOS: 15 or 18 and SEQ ID NOS: 16 or 19.

20 18. The method of claim 13, wherein the nucleic acid amplification reaction comprises a nested polymerase chain reaction.

19. The method of claim 13, wherein at least one of the primer nucleic acids comprises at least one label.

25 20. The method of claim 19, wherein (b) comprises detecting a detectable signal produced by the label, or amplifying a detectable signal produced by the label to produce an amplified signal and detecting the amplified signal.

21. The method of claim 13, wherein (b) comprises monitoring binding between the amplicons and at least one oligonucleotide having a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID

30

NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant.

22. The method of claim 21, wherein the oligonucleotide comprises at least one label and/or at least one quencher moiety.

5 23. The method of claim 22, wherein the oligonucleotide comprises a 5'-nuclease probe having a sequence selected from SEQ ID NO: 27 or SEQ ID NO: 28.

24. A method of determining a presence of a severe acute respiratory syndrome coronavirus in a sample, the method comprising:

10 (a) contacting nucleic acids and/or amplicons thereof from the sample with one or more oligonucleotides that comprise at least one nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant; and,

15 (b) monitoring binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides, wherein detectable binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides, determines the presence of the severe acute respiratory syndrome coronavirus in the sample.

20 25. The method of claim 24, wherein (a) comprises contacting the nucleic acids and/or amplicons thereof with the oligonucleotides in solution at a temperature of at least 42°C for at least 15 minutes, wherein a total weight of the solution comprises about 50% formalin and comprises heparin at a concentration of about 1 mg/ml.

25 26. The method of claim 24, comprising repeating (a) and (b) at least once using at least one additional sample and comparing the binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides, of (b) with at least one repeated (b).

30 27. The method of claim 24, wherein at least one segment of the nucleic acids is amplified prior to or during (a) using at least one nucleic acid amplification technique to produce the amplicons and (b) comprises monitoring the binding between the nucleic acids and/or amplicons thereof, and the oligonucleotides, during or after amplification.

28. A composition comprising a sample derived from a subject and at least one oligonucleotide that comprises a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 5 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, which oligonucleotide consists of 100 or fewer nucleotides.

29. A kit, comprising:

(a) at least one oligonucleotide that comprises a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a 10 substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, which oligonucleotide consists of 100 or fewer nucleotides; and one or more of:

(b) instructions for determining a presence of a severe acute respiratory 15 syndrome coronavirus in a sample by monitoring binding between nucleic acids and/or amplicons thereof from the sample and the oligonucleotide, wherein the presence of the severe acute respiratory syndrome coronavirus in the sample is unknown or unsubstantiated; or,

(c) at least one container for packaging at least the oligonucleotide.

20 30. The kit of claim 29, further comprising at least one enzyme.

31. A system for detecting a severe acute respiratory syndrome coronavirus in a sample, comprising:

(a) at least one oligonucleotide that comprises a nucleic acid with a sequence selected from the group consisting of: SEQ ID NOS: 1-12 and 15-24, a 25 substantially identical variant thereof wherein the variant has at least 90% sequence identity to one of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant, which oligonucleotide consists of 100 or fewer nucleotides;

(b) at least one detector that detects binding between nucleic acids 30 and/or amplicons thereof from the sample and the oligonucleotide; and,

(c) at least one controller operably connected to the detector, which controller comprises one or more instructions sets that correlate the binding detected by the detector with a presence of the severe acute respiratory syndrome coronavirus in the sample.

5 32. A system, comprising:

(a) computer or computer readable medium comprising a data set that comprises a plurality of character strings that correspond to a plurality of sequences that correspond to one or more of: SEQ ID NOS: 1-12 and 15-24, a substantially identical variant thereof wherein the variant has at least 90% sequence identity to one 10 of SEQ ID NOS: 1-12 and 15-24, and complements of SEQ ID NOS: 1-12 and 15-24 and the variant; and,

(b) an automatic synthesizer coupled to an output of the computer or computer readable medium, which automatic synthesizer accepts instructions from the computer or computer readable medium, which instructions direct synthesis of one or 15 more nucleic acids that correspond to one or more character strings in the data set.

1/4

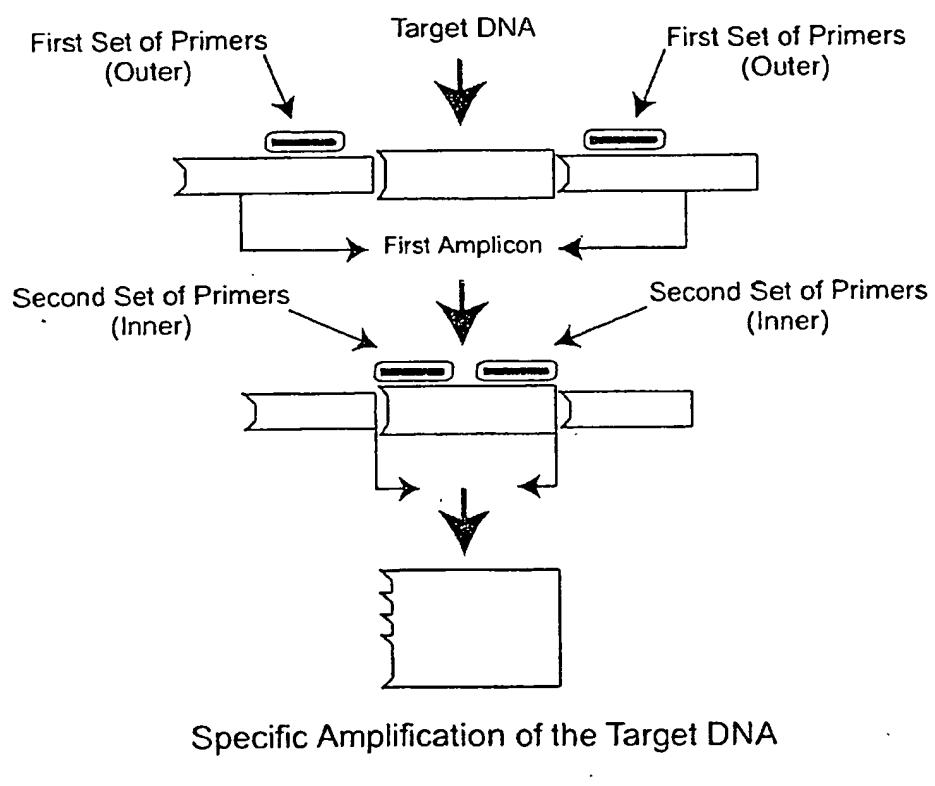


Fig. 1

JC20 Rec'd PCT/PTO 07 OCT 2005

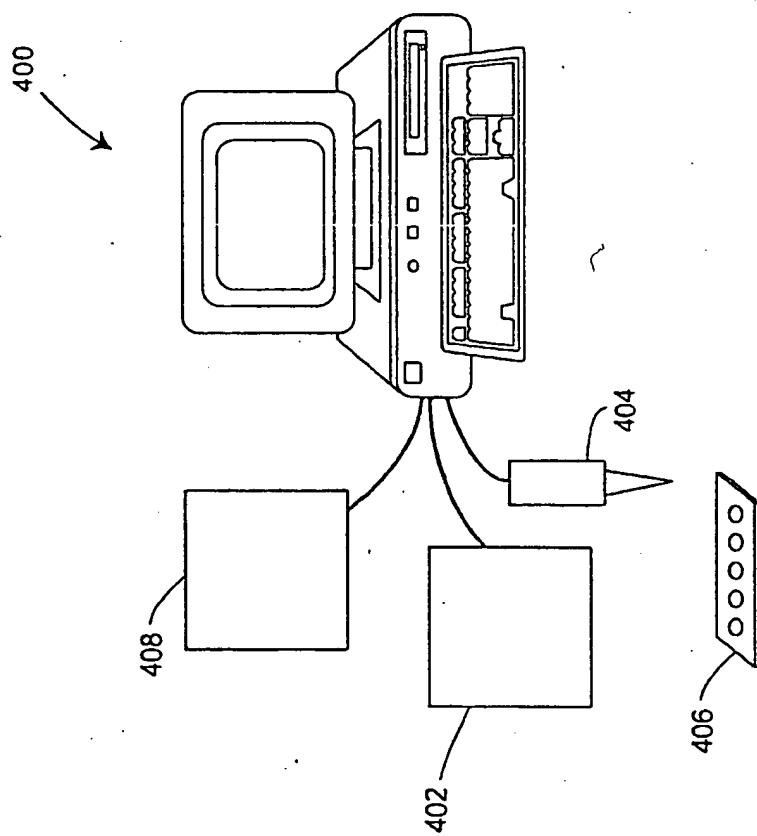


Fig. 2

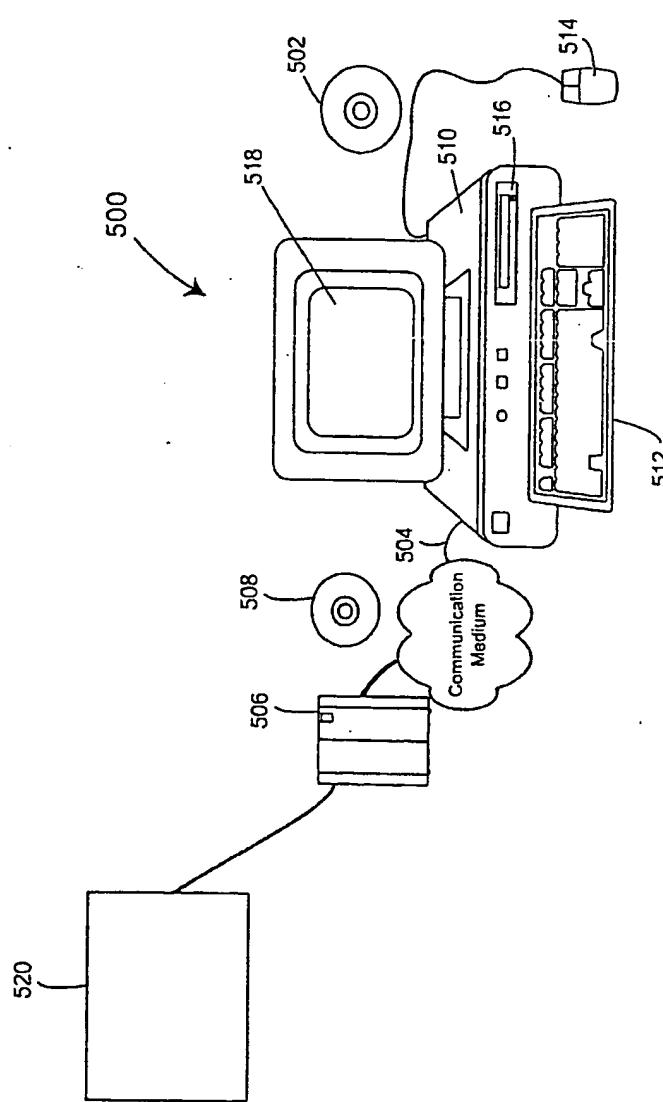


Fig. 3

4/4

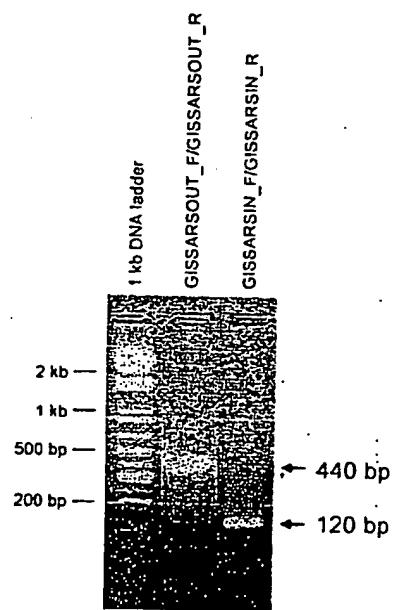


Fig. 4

The entire nucleotide sequence of SARS-CoV (Urbani strain).  
The genome is 29, 727 nucleotides in length from 5' leader to 3'end.

1 TTATTAGGTT TTTACCTACC CAGGAAAAGC CAACCAACCT CGATCTCTTG  
51 TAGATCTGTT CTCTAACGA ACTTTAAAAT CTGTGTAGCT GTCGCTCGGC  
101 TGCATGCCTA GTGCACCTAC GCAGTATAAA CAATAATAAA TTTTACTGTC  
151 GTTGACAAGA AACGAGTAAC TCGTCCCTCT TCTGCAGACT GCTTACGGTT  
201 TCGTCCGTGT TGCAGTCGAT CATCAGCATA CCTAGGTTTC GTCCGGGTGT  
251 GACCGAAAGG TAAGATGGAG AGCCTGTTTC TTGGTGTCAA CGAGAAAACA  
301 CACGTCCAAC TCAGTTGCC TGTCCCTCAG GTTAGAGACG TGCTAGTGC  
351 TGGCTTCGGG GACTCTGTGG AAGAGGCCCT ATCGGAGGCA CGTGAACACC  
401 TCAAAAATGG CACTTGTGGT CTAGTAGAGC TGAAAAGG CGTACTGCC  
451 CAGCTTGAAC AGCCCTATGT GTTCATTAAC CGTTCTGATG CCTTAAGCAC  
501 CAATCACGGC CACAAGGTCG TTGAGCTGGT TGCAGAAATG GACGGCATT  
551 AGTACGGTCG TAGCGGTATA ACACTGGGAG TACTCGTGC ACATGTGGC  
601 GAAACCCCAA TTGCATACCG CAATGTTCTT CTTCGTAAGA ACGGTAA  
651 GGGAGCCGGT GGTCACTAGCT ATGGCATCGA TCTAAAGTCT TATGACTTAG  
701 GTGACGAGCT TGGCACTGAT CCCATTGAAAG ATTATGAACA AACTGGAAC  
751 ACTAACGATG GCAGTGGTGC ACTCCGTGAA CTCACACTGT AGCTCAATGG  
801 AGGTGCAGTC ACTCGCTATG TCGACAAACAA TTCTGTGGC CCAGATGGGT  
851 ACCCTCTTGA TTGCACTAAA GATTTCTCG CACCGCGGGG CAAGTCATG  
901 TGCACCTTT CCGAACAACT TGATTACATC GAGTCGAAGA GAGGTGTCTA  
951 CTGCTGCCGT GACCATGAGC ATGAAATTGC CTGGTTCACT GAGGGCTCTG  
1001 ATAAGAGCTA CGAGCACCAAG ACACCCCTCG AAATTAAAGAG TGCCAAGAAA  
1051 TTGACACTT TCAAAAGGGGA ATGCCAAAG TTGTGTTTC CTCTTAAC  
1101 AAAAGTCAAA GTCACTAAC CACGTGTTGA AAAAGAAAAG ACTGAGGGTT  
1151 TCATGGGGCG TATACGCTCT GTGTACCCCTG TTGCATCTCC ACAGGGAGTGT  
1201 AACAAATATGC ACTTGTCTAC CTTGATGAAA TGTAACTCATT GCGATGAAGT  
1251 TTGATGGCG ACGTGCGACT TTCTGAAAGC CACTTGTGAA CATTGTGGCA  
1301 CTGAAAATT AGTTATTGAA GGACCTACTA CATGTGGGT CCTACCTACT  
1351 AATGCTGTAG TGAATGCGC ATGTCCTGCC TGTCAAGACC CAGAGATTGG  
1401 ACCTGAGCAT AGTGTGCGAG ATTATCACAA CCACTCAAAC ATTGAAACTC  
1451 GACTCCGCAA GGGAGGTTAGG ACTAGATGTT TTGGAGGCTG TGTGTTGCC  
1501 TATGTTGGCT GCTATAATAA GCGTGCCTAC TGGGTTCCCTC GTGCTAGTGC  
1551 TGATATTGGC TCAGGCCATA CTGGCATTAC TGGTGAACAT GTGGAGACCT  
1601 TGAATGAGGA TCTCCTTGAG ATACTGAGTC GTGAACGTGT TAACATTAAC  
1651 ATTGTTGGCG ATTTCATTGAATGAAGAG GTTGCATCA TTTGGC  
1701 TTTCTCTGCT TCTACAAGTG CCTTTATTGA CACTATAAG AGTCTTGATT  
1751 ACAAGTCTTT CAAAACCATT GTTGAGTCCT GCGGTAAC TAAAGTTAC  
1801 AAGGGAAAGC CCGTAAAAGG TGCTTGGAAC ATTGGACAAC AGAGATCAGT  
1851 TTAAACACCA CTGTGTGGTT TTCCCTCACA GGCTGCTGGT GTTATCAGAT  
1901 CAATTTTGC GCGCACACTT GATGCAGCAA ACCACTCAAT TCCTGATT  
1951 CAAAGAGCAG CTGTCACCAT ACTTGATGGT ATTTCATGAAAGTCA  
2001 TCTTGTGAC GCCATGGTTT ATACTTCAGA CCTGCTCACC AACAGTGTCA  
2051 TTATTATGGC ATATGTAACG GGTGGTCTTG TACAACAGAC TTCTCAGTGG  
2101 TTGTCATAATC TTTGGGCAC TACTGTTGAA AAACCTCAGGC CTATCTTGA  
2151 ATGGATTGAG GCGAAACTTA GTGCAGGAGT TGAATTCTC AAGGATGCTT  
2201 GGGAGATTCT CAAATTCTC ATTACAGGTG TTTTGACAT CGTCAAGGGT  
2251 CAAATACAGG TTGCTTCAGA TAACATCAAG GATTGTGTAA AATGCTTCAT  
2301 TGATGTTGTT ACAAGGCAC TCGAAATGTG CATTGATCAA GTCACTATCG  
2351 CTGGCGCAA GTTGCATCA CTCAACTTAG GTGAAGTCTT CATCGCTCAA  
2401 AGCAAGGGAC TTACCGTCA GTGTATACGT GGCAAGGAGC AGCTGCAACT  
2451 ACTCATGCCT CTTAAGGCAC CAAAAGAAGT AACCTTCTT GAAGGTGATT

2501 CACATGACAC AGTACTTACC TCTGAGGAGG TTGTTCTCAA GAACGGTGAA  
2551 CTCGAAGCAC TCGAGACGCC CGTTGATAGC TTCACAAATG GAGCTATCGT  
2601 TGGCACACCA GTCTGTGAA ATGGCCTCAT GCTCTTAGAG ATTAAGGACA  
2651 AAGAACATA CTGCGCATTG TCTCCTGGTT TACTGGCTAC AAACAATGTC  
2701 TTTCGCTTAA AAGGGGGTGC ACCAATTAAA GGTGTAACCT TTGGAGAAGA  
2751 TACTGTTGG GAAGTTCAAG GTTACAAGAA TGTGAGAATC ACATTGAGC  
2801 TTGATGAACG TGTGACAAA GTGCTTAATG AAAAGTGCTC TGTCTACACT  
2851 GTTGAATCCG GTACCGAAGT TACTGAGTTT GCATGTGTTG TAGCAGAGGC  
2901 TGTTGTGAAG ACTTTACAAC CAGTTCTGA TCTCCTTACCAACATGGTA  
2951 TTGATCTTGA TGAGTGGAGT GTAGCTACAT TCTACTTATT TGATGATGCT  
3001 GGTGAAGAAA ACTTTCATC ACGTATGTAT TGTTCTTACCAACCTCCAGA  
3051 TGAGGAAGAA GAGGACGATG CAGAGTGTGA GGAAGAAGAA ATTGATGAAA  
3101 CCTGTGAACA TGAGTACGGT ACAGAGGATG ATTATCAAGG TCTCCTCTG  
3151 GAATTTGGTG CCTCAGCTGA AACAGTTCGA GTTGGAGGAAG AAGAAGAGGA  
3201 AGACTGGCTG GATGATACTA CTGAGCAATC AGAGATTGAG CCAGAACAG  
3251 AACCTACACC TGAAGAACCA GTTAATCACT TTACTGGTTA TTAAAACCTT  
3301 ACTGACAATG TTGCCATTAA ATGTGTTGAC ATCGTTAAGG AGGCACAAAG  
3351 TGCTAATCCT ATGGTGTGATTG TAAATGCTGC TAACATACAC CTGAAACATG  
3401 GTGGTGGTGT AGCAGGTTGCA CTCAACAAAGG CAACCAATGG TGCCATGCAA  
3451 AAGGAGAGTG ATGATTACAT TAAGCTAAAT GGCCCTCTTA CAGTAGGAGG  
3501 GTCTTGTGTTG CTTCTGGAC ATAATCTTC TAAGAAGTGT CTGCATGTTG  
3551 TTGGACCTAA CCTAAATGCA GGTGAGGACA TCCAGCTCT TAAGGCAGCA  
3601 TATGAAAATT TCAATTACA GGACATCTTA CTTGCACCAT TGTTGTCAGC  
3651 AGGCATATTG GTGCTAACAC CACTTCAGTC TTACAAGTG TGCGTGCAGA  
3701 CGGTTCTGAC ACAGGTTAT ATTGCAGTCA ATGACAAAGC TCTTATGAG  
3751 CAGGTTGTCA TGGATTATCT TGATAACCTG AAGCCTAGAG TGGAAGCACC  
3801 TAAACAAGAG GAGCCACCAA ACACAGAAGA TTCCAAAAGT GAGGAGAAAT  
3851 CTGTCGTACA GAAGCCTGTC GATGTGAAGC CAAAAATTAA GGCCCTGCATT  
3901 GATGAGGTTA CCACAAACACT GGAAGAAACT AAGTTCTTA CCAATAAGTT  
3951 ACTCTTGTGTT GCTGATATCA ATGGTAAGCT TTACCATGAT TCTCAGAACAA  
4001 TGCTTAGAGG TGAAGATATG TCTTCTTG AGAAGGATGC ACCTTACATG  
4051 GTAGGTGATG TTATCACTAG TGGTGTATC ACTTGTGTTG TAATACCCTC  
4101 CAAAAAGGCT GGTGGCACTA CTGAGATGCT CTCAGAGCT TTGAAGAAAG  
4151 TGCCAGTTGA TGAGTATATA ACCACGTACC CTGGACAAGG ATGTGCTGGT  
4201 TATACACTTG AGGAAGCTAA GACTGCTCTT AAGAAATGCA AATCTGCATT  
4251 TTATGTA CTTCTGAGAAG CACCTAATGC TAAGGAAGAG ATTCTAGGAA  
4301 CTGTATCCTG GAATTGAGA GAAATGCTTG CTCATGCTGA AGAGACAAGA  
4351 AAATTAATGC CTATATGCAT GGATGTTAGA GCCATAATGG CAACCATCCA  
4401 ACGTAAGTAT AAAGGAATTAA AAATTCAGA GGGCATCGTT GACTATGGTG  
4451 TCCGATTCTT CTTTATACT AGTAAAGAGC CTGTAGCTTC TATTATTACG  
4501 AAGCTGAACCTCTAAATGAA GCGCCTGTC ACAATGCCAA TTGGTTATGT  
4551 GACACATGGT TTTAATCTTG AAGAGGCTGC GCGCTGTATG CGTTCTCTTA  
4601 AAGCTCTGC CGTAGTGTCA GTATCATCAC CAGATGCTGT TACTACATAT  
4651 AATGGATACC TCACTTCGTC ATCAAAGACA TCTGAGGAGC ACTTTGTAGA  
4701 AACAGTTCT TTGGCTGGCT CTTACAGAGA TTGGTCTTAT TCAGGACAGC  
4751 GTACAGAGTT AGGTGTTGAA TTTCTTAAGC GTGGTGACAA AATTGTGTAC  
4801 CACACTCTGG AGAGCCCCGT CGAGTTCTAT CTTGACGGTG AGGTTCTTTC  
4851 ACTTGACAAA CAAAGAGTC TCTTATCCCT GCGGGAGGTT AAGACTATAA  
4901 AAGTGTTCAC AACTGTGGAC AACACTAATC TCCACACACA GCTTGTGGAT  
4951 ATGTCTATGA CATATGGACA GCAGTTGGT CCAACATACT TGGATGGTGC  
5001 TGATGTTACA AAAATTAAC CTCATGTTAA TCATGAGGGT AAGACTTTCT  
5051 TTGTA CTTGACTTACGAC AACTACGTA GTGAAGCTT CGAGTACTAC  
5101 CATACTCTTG ATGAGAGTTT TCTTGGTAGG TACATGTCTG CTTAAACCA

5151 CACAAAGAAA TGGAAATTC CTCAGTTGG TGGTTAACT TCAATTAAT  
5201 GGGCTGATAA CAATTGTTAT TTGCTAGTG TTTTATTAGC ACTTCAACAG  
5251 CTTGAAGTCA AATTCAATGC ACCAGCACCT CAAGAGGCCT ATTATAGAGC  
5301 CCGTGCTGGT GATGCTGCTA ACTTTGTGC ACTCATACTC GCTTACAGTA  
5351 ATAAAACGTG TGGCGAGCTT GGTGATGTCA GAGAAACTAT GACCCATCTT  
5401 CTACAGCATG CTAATTGGA ATCTGCAAAG CGAGTTCTTA ATGTGGTGTG  
5451 TAAACATTGT GGTAGAAAAA CTACTACCTT AACGGGTGTA GAAGCTGTGA  
5501 TGTATATGGG TACTCTATCT TATGATAATC TTAAGACAGG TGTTCCATT  
5551 CCATGTGTGT GTGGCTGTGA TGCTACACAA TATCTAGTAC ACAAGAGTC  
5601 TTCTTTGTT ATGATGTCTG CACCACCTGC TGAGTATAAA TTACAGCAAG  
5651 GTACATTCTT ATGTGCGAAT GAGTACACTG GTAATATCA GTGTGGTCAT  
5701 TACACTCATA TAATGCTAA GGAGACCCTC TATCGTATTG ACGGAGCTCA  
5751 CCTTACAAAG ATGTCAGAGT ACAAAGGACC AGTGAATGAT GTTTCTACA  
5801 AGGAAACATC TTACACTACA ACCATCAAGC CTGTGTCGTA TAAACTCGAT  
5851 GGAGTTACTT ACACAGAGAT TGAACCAAAA TTGGATGGGT ATTATAAAA  
5901 GGATAATGCT TACTATACAG AGCAGCCTAT AGACCTTGTA CCAACTCAAC  
5951 CATTACCAAA TGCAGTTT GATAATTCA AACTCACATG TTCTAACACA  
6001 AAATTTGCTG ATGATTTAAA TCAAATGACA GGCTTCACAA AGCCAGCTTC  
6051 ACGAGAGCTA TCTGTCACAT TCTTCCAGA CTTGAATGGC GATGTAGTGG  
6101 CTATTGACTA TAGACACTAT TCAGCGAGTT TCAAGAAAGG TGCTAAATTA  
6151 CTGCATAAGC CAATTGTTTG GCACATTAAC CAGGCTACAA CCAAGACAAC  
6201 GTTCAACCA AACACTTGGT GTTACGTTG TCTTGGAGT ACAAAGCCAG  
6251 TAGATACTTC AAATTCAATTGAAAGTTCTGG CAGTAGAAGAACACAAAGGA  
6301 ATGGACAATC TTGCTTGTGA AAGTCAACAA CCCACCTCTG AAGAAGTAGT  
6351 GGAAAATCCT ACCATACAGA AGGAAGTCAT AGAGTGTGAC GTGAAAAC  
6401 CCGAAGTTGT AGGAATGTC ATACTTAAAC CATCAGATGA AGGTGTTAAA  
6451 GTAACACAAG AGTTAGGTCA TGAGGATCTT ATGGCTGCTT ATGTGGAAAA  
6501 CACAAGCATT ACCATTAAGAACCTAAATGCA GCTTCACTA GCCTTAGGTT  
6551 TAAAAAACAT TGCCACTCAT GGTATTGCTG CAATTAATAG TGTCCTTGG  
6601 AGTAAAATTT TGCTTATGT CAAACCAATTGCA TTAGGACAAG CAGCAATTAC  
6651 AACATCAAAT TGCGCTAAGA GATTAGCACA ACGTGTGTTT AACAAATTATA  
6701 TGCCTTATGT GTTACATTA TTGTTCCAAT TGTGTACTTT TACTAAAAGT  
6751 ACCAATTCTA GAATTAGAGC TTCACACTT ACAACTATTG CTAAAATAG  
6801 TGTTAAGAGT GTGCTAAAT TATGTTGGA TGCCGGCATT AATTATGTGA  
6851 AGTCACCCAA ATTTCTAAA TTGTTACAA TCGCTATGT GCTATTGTTG  
6901 TTAAGTATTT GCTTAGGTTC TCTAATCTGT GTAATGCTG CTTTGGTGT  
6951 ACTCTTATCT AATTTGGTG CTCCTTCATA TTGTAATGGC GTTAGAGAAAT  
7001 TGTATCTTAA TTGCTCTAAC GTTACTACTA TGAGTTCTG TGAAGGTTCT  
7051 TTTCTTGCA GCATTGTTT AAGTGGATTAA GACTCCCTTG ATTCTTATCC  
7101 AGCTCTTGAA ACCATTCAAG TGACGATTTC ATCGTACAAG CTAGACTTGA  
7151 CAATTAGG TCTGGCCGCT GAGTGGGTTT TGGCATATAT GTTGTTCACA  
7201 AAATTCTTT ATTATTTAGG TCTTCAGCT ATAATGCAGG TGTTCTTGG  
7251 CTATTTGCT AGTCATTCTA TCAGCAATTG TTGGCTCATG TGGTTATCA  
7301 TTAGTATTGT ACAAAATGGCA CCCGTTCTG CAATGGTTAG GATGTACATC  
7351 TTCTTGCTT CTTCTACTA CATATGGAAG AGCTATGTTCA ATATCATGGA  
7401 TGGTGCACC TCTTCGACTT GCATGATGTG CTATAAGCGC AATCGTCCA  
7451 CACGCGTTGA GTGTACAACCTT ATTGTTAATG GCATGAAGAG ATCTTCTAT  
7501 GTCTATGCAA ATGGAGGCCG TGGCTTCTGC AAGACTCACA ATTGGAATTG  
7551 TCTCAATTGT GACACATTTC GCACTGGTAG TACATTCTT AGTGTGAAAG  
7601 TTGCTCGTGA TTGTCACTC CAGTTAAAAA GACCAATCAA CCCTACTGAC  
7651 CAGTCATCGT ATATTGTTGA TAGTGTGCT GTGAAAATG GCGCCCTCA  
7701 CCTCTACTTT GACAAGGCTG GTCAAAAGAC CTATGAGAGA CATCCGCTCT  
7751 CCCATTGTG CAATTAGAC AATTGAGAG CTAACAAACAC TAAAGGTTCA

7801 CTGCCTATT AATGTCATAGT TTTTGATGGC AAGTCCAAAT GCGACGGAGTC  
7851 TGCTTCTAAG TCTGCTTCTG TGTACTACAG TCAGCTGATG TGCCAACCTA  
7901 TTCTGTTGCT TGACCAAGTT CTTGTATCAG ACGTTGGAGA TAGTACTGAA  
7951 GTTCCCGTTA AGATGTTGA TGCTTATGTC GACACCTTT CAGCAACTT  
8001 TAGTGTTCCT ATGGAAAAAC TTAAGGCAC TGTGCTACA GCTCACAGCG  
8051 AGTTAGCAAA GGGTGTAGCT TTAGATGGTG TCCTTCTAC ATTCGTGTCA  
8101 GCTGCCGAC AAGGTGTTGT TGATACCGAT GTTGACACAA AGGATGTTAT  
8151 TGAATGTCTC AAACTTTACATCCTGA CTTAGAAGTG ACAGGTGACA  
8201 GTTGTAAACAA TTTCATGTC ACCTATAATA AGGTTGAAAA CATGACGCC  
8251 AGAGATCTG GCGCATGTAT TGACTGTAAT GCAAGGCATA TCAATGCC  
8301 AGTAGCAAAA AGTCACAATG TTTCACTCAT CTGGAATGTA AAAGACTACA  
8351 TGTCTTATC TGAAACAGCTG CGTAAACAAA TTCTGTTGC TGCCAGAAG  
8401 AACAAACATAC CTTTAGACT AACTTGCT ACAACTAGAC AGGTTGCAA  
8451 TGTCTAACT ACTAAAATCT CACTCAAGGG TGTTAAGATT GTTAGTACTT  
8501 GTTTAAACT TATGCTTAAG GCCACATTAT TGTCGTTCT TGCTGCATTG  
8551 GTTTGTTATA CGTTATGCC AGTACATACA TTGTCATCC ATGATGGTTA  
8601 CACAAATGAA ATCATTGGTT ACAAAAGCCAT TCAGGATGGT GTCACTCGTG  
8651 ACATCATTTC TACTGATGAT TGTTTGCAA ATAAACATGC TGGTTTGAC  
8701 GCATGGTTA GCCAGCGTGG TGGTCATAC AAAATGACA AAAGCTGCC  
8751 TGTAGTAGCT GCTATCATTAA CAAGAGAGAT TGTTTCATA GTGCCTGGCT  
8801 TACCGGGTAC TGTGCTGAGA GCAATCAATG GTGACTTCCT GCATTTCTA  
8851 CCTCGTGTCTT TTAGTGTGT TGGCAACATT TGCTACACAC CTTCCAAACT  
8901 CATTGAGTAT AGTGATTTG CTACCTCTGC TTGCGTTCT GCTGCTGAGT  
8951 GTACAAATTTT TAAGGATGCT ATGGGCAAC CTGTGCCATA TTGTTATGAC  
9001 ACTAATTGCTAGAGGGTTC TATTCCTTAT AGTGAGCTTC GTCCAGACAC  
9051 TCGTTATGTG CTTATGGATG GTTCCATCAT ACAGTTTCCT AACACTTAC  
9101 TGGAGGGTTC TGTTAGAGTA GTAACAACTT TTGATGCTGA GTACTGTAGA  
9151 CATGGTACAT GCGAAAGGTC AGAAGTAGGT ATTTGCCTAT CTACCAAGTGG  
9201 TAGATGGGTT CTTAATAATG AGCATTACAG AGCTCTATCA GGAGTTTCT  
9251 GTGGTGTGTA TGCATGAAT CTCTAGCTA ACATCTTAC TCCTCTTG  
9301 CAACCTGTGG GTGCTTCTAGA TGTCGTTCT TGAGTAGTGG CTGGTGGTAT  
9351 TATTGCCATA TTGGTGACTT GTGCTGCCATA CTACTTTATG AAATTAGAC  
9401 GTGTTTTGG TGAGTACAAC CATGTTGTTG CTGCTAATGC ACTTTGTT  
9451 TTGATGTCTT TCACTATACT CTGCTGGTA CCAGCTTACA GCTTTCTGCC  
9501 GGGAGTCTAC TCAGCTTTT ACTTGACTT GACATTCTAT TTCAACATG  
9551 ATGTTTCATT CTTGGCTCAC CTTCAATGGT TTGCCATGTT TTCTCTTATT  
9601 GTGCCCTTTT GGATAACAGC AATCTATGTA TTCTGTATTT CTCTGAAGCA  
9651 CTGCCATTGG TTCTTTAACAA ACTATCTTAG GAAAAGAGTC ATGTTAATG  
9701 GAGTTACATT TAGTACCTTC GAGGAGGCTG CTTGTGTAC CTTTTGCTC  
9751 AACAAAGGAA TGACCTAAA ATTGCGTAGC GAGACACTGT TGCCACTTAC  
9801 ACAGTATAAC AGGTATCTT CTCTATATAA CAAGTACAAG TATTCACTG  
9851 GAGCCTTACA TACTACAGC TATCGTGAAG CAGCTTGTG CCACCTAGCA  
9901 AAGGCTCTAA ATGACTTTAG CAACTCAGGT GCTGATGTT CTCACCAACC  
9951 ACCACAGACA TCAATCACCTT CTGCTGTTCT GCAGAGTGGT TTTAGGAAAA  
10001 TGGCATTCCC GTCAGGCAAA GTTGAAGGGT GCATGGTACA AGTAACCTGT  
10051 GGAACATACAA CTCTTAATGG ATTGTGGTTG GATGACACAG TATACTGTCC  
10101 AAGACATGTC ATTGACAG CAGAAGACAT GCTTAATCCT AACTATGAAG  
10151 ATCTGCTCAT TCGCAAATCC AACCATAGCT TTCTGTGTC GGCTGGCAAT  
10201 GTTCAACTTC GTGTTATTGG CCATTCTATG CAAAATTGTC TGCTTAGGCT  
10251 TAAAGTTGAT ACTTCTAACCT CAAACACCC CAAAGTATAAA TTTGTCCGTA  
10301 TCCAACCTGG TCAAACATT TCAAGACACC CAAAGTATAAA TTTGTCCGTA  
10351 TCTGGTGTGTT ATCAGTGTGC CATGAGACCT AATCATACCA TAAAGGTT  
10401 TTTCCTTAAT GGATCATGTG GTAGTGTGG TTAAACATT GATTATGATT

10451 GCGTGTCTT CTGCTATATG CATCATATGG AGCTTCCAAC AGGAGTACAC  
 10501 GCTGGTACTG ACTTAGAAGG TAAATTCTAT GGTCCATTG TTGACAGACA  
 10551 AACTGCACAG GCTGCAGGTA CAGACACAAC CATAACATTA AATGTTTGG  
 10601 CATGGCTGTA TGCTGCTGTT ATCAATGGTG ATAGGTGGTT TCTTAATAGA  
 10651 TTCACCACTA CTTTGAATGA CTTAACCTT GTGGCAATGA AGTACAACTA  
 10701 TGAACCTTG ACACAAGATC ATGTTGACAT ATTGGGACCT CTTCTGCTC  
 10751 AAACAGGAAT TGCCGTCTTA GATATGTGTG CTGCTTGAA AGAGCTGCTG  
 10801 CAGAATGGTA TGAATGGTCG TACTATCCTT GGTAGCACTA TTTTAAAGA  
 10851 TGAGTTTACA CCATTTGATG TTGTTAGACA ATGCTCTGGT GTTACCTTCC  
 10901 AAGGTAAGTT CAAGAAAATT GTTAAGGGCA CTCATCATTG GATGCTTTA  
 10951 ACTTTCTTGA CATCACTATT GATTCTTGT CAAAGTACAC AGTGGTCACT  
 11001 GTTTTCTT GTTACGAGA ATGCTTCTT GCCATTTACT CTTGGTATTAA  
 11051 TGGCAATTGC TGCATGTGCT ATGCTGCTTG TTAAGCATAA GCACGCATT  
 11101 TTGTGCTTGT TTCTGTTACC TTCTCTTGCA ACAGTTGCTT ACTTTAATAT  
 11151 GGTCTACATG CCTGCTAGCT GGGTGATGCG TATCATGACA TGGCTTGAAT  
 11201 TGGCTGACAC TAGCTTGTCT GGTTATAGGC TTAAGGATTG TGTTATGTAT  
 11251 GCTTCAGCTT TAGTTTGTCT TATTCTCATG ACAGCTCGCA CTGTTTATGA  
 11301 TGATGCTGCT AGACGTGTTT GGACACTGAT GAATGTCATT ACACTTGTTT  
 11351 ACAAAAGTCTA CTATGGTAAT GCTTAGATC AAGCTATTC CATGTGGGCC  
 11401 TTAGTTATTCTGTACACCTC TAACTATTCTT GGTGTCGTTA CGACTATCAT  
 11451 GTTTTTAGCT AGAGCTATAG TGTTTGTGTG TGTTGAGTAT TACCCATTGT  
 11501 TATTTATTAC TGGCAACACC TTACAGTGTA TCATGCTTGT TTATTGTTTC  
 11551 TTAGGCTATT GTTGCTGCTG CTACTTTGGC CTTTTCTGTT TACTCAACCG  
 11601 TTACTTCAGG CTTACTCTTGT GTGTTATGA CTACTGGTC TCTACACAAG  
 11651 AATTAGGTA TATGAACCTCC CAGGGGCTTT TGCCCTCTAA GAGTAGTATT  
 11701 GATGCTTCA AGCTTAACAT TAAGTTGTG TGTTGAGTGGAG GTAAACCATG  
 11751 TATCAAGGTT GCTACTGTAC AGTCTAAAAT GTCTGACGTA AAGTGCACAT  
 11801 CTGTGGTACT GCTCTCGTGT CTTAACACAC TTAGAGTAGA GTCATCTTCT  
 11851 AAATTGTGGG CACAATGTGT ACAACTCCAC AATGATATTCT TTCTTGCAAA  
 11901 AGACACAACT GAAGCTTCG AGAAGATGGT TTCTCTTTG TCTGTTTGC  
 11951 TATCCATGCA GGGTGCTGTA GACATTAATA GGTTGTGCGA GGAAATGCTC  
 12001 GATAACCGTG CTACTCTTCA GGCTATTGCT TCAGAATTAA GTTCTTACC  
 12051 ATCATATGCC GCTTATGCCA CTGGCCAGGA GGCTATGAG CAGGCTGTAG  
 12101 CTAATGGTGA TTCTGAAGTC GTTCTCAAAA AGTTAAAGAA ATCTTGAAT  
 12151 GTGGCTTAAT CTGAGTTTGA CCGTGATGCT GCCATGCAAC GCAAGTTGGA  
 12201 AAAGATGGCA GATCAGGCTA TGACCCAAAT GTACAAACAG GCAAGATCTG  
 12251 AGGACAAGAG GGCAAAAGTA ACTAGTGCTA TGCAAACAAT GCTCTTCACT  
 12301 ATGCTTAGGA AGCTTGATAA TGATGCACTT AAAACATTA TCAACAATGC  
 12351 GCGTGATGGT TGTTGTCAC TCAACATCAT ACCATTGACT ACAGCAGCCA  
 12401 AACTCATGGT TGTTGTCCT GATTATGGTA CCTACAAGAA CACTTGTGAT  
 12451 GGTAAACACCT TTACATATGC ATCTGCACTC TGGGAAATCC AGCAAGTTGT  
 12501 TGATGCGGAT AGCAAGATTG TTCAACTTAG TGAAATTAAC ATGGACAATT  
 12551 CACCAAATTG GGCTGGCCT CTTATTGTTA CAGCTCTAAAG AGCCAACCTCA  
 12601 GCTGTTAAC TACAGAATAA TGAACGTAGT CCAGTAGCAC TACCGACAGAT  
 12651 GTCCTGTGCG GCTGGTACCA CACAAACAGC TTGTAATGAT GACAATGCA  
 12701 TTGCCTACTA TAACAATTG AAGGGAGGTA GGTTTGTGCT GGCATTACTA  
 12751 TCAGACCACC AAGATCTCAA ATGGGCTAGA TTCCCTAAGA GTGATGGTAC  
 12801 AGGTACAATT TACACAGAAC TGGAACCAAC TTGTAGGTTT GTTACAGACA  
 12851 CACCAAAGG GCCTAAAGTG AAATACCTGT ACTTCATCAA AGGCTTAAAC  
 12901 AACCTAAATA GAGGTATGGT GCTGGGCAGT TTAGCTGCTA CAGTACGTCT  
 12951 TCAGGCTGGA AATGCTACAG AAGTACCTGC CAATTCAACT GTGCTTTCCT  
 13001 TCTGTGCTTT TGCAAGTAGAC CCTGCTAAAG CATATAAGGA TTACCTAGCA  
 13051 AGTGGAGGAC AACCAATCAC CAACTGTGTG AAGATGTTGT GTACACACAC

13101 TGGTACAGGA CAGGCAATT A CTGTAACACC AGAAGCTAAC ATGGACCAAG  
 13151 AGTCCTTGG TGGTGCCTCA TGTTGCTGT ATTGTAGATG CCACATTGAC  
 13201 CATCCAAATC CAAAGGATT CTGTGACTG AAAGGTAAGT ACGTCCAAAT  
 13251 ACCTACCACT TGTGCTAATG ACCCAGTGGG TTTTACACTT AGAAACACAG  
 13301 TCTGTACCGT CTGCGGAATG TGGAAAGGTT ATGGCTGTAG TTGTGACCAA  
 13351 CTCCGCGAAC CCTTGATGCA GTCTGCGGAT GCATCAACGT TTTAAACGG  
 13401 GTTTGCGGTG TAAGTGCAGC CCGTCTTACA CGTGCAGCA CAGGCACTAG  
 13451 TACTGATGTC GTCTACAGGG CTTTGATAT TTACAAACGAA AAAGTTGCTG  
 13501 GTTTGCAAA GTTCTAAACTAATGCT GTCGCTTCCA GGAGAAGGAT  
 13551 GAGGAAGGCA ATTATTAGA CTCTTACTTT GTAGTTAGA GGCATACTAT  
 13601 GTCTAACTAC CAACATGAAG AGACTATTAA TAACTTGGTT AAAGATTGTC  
 13651 CAGCGGTTGC TGTCATGAC TTTTCAAGT TTAGAGTAA TGTTGACATG  
 13701 GTACACACATA TATCACGTC GCGCTAACT AAATACACAA TGGCTGATT  
 13751 AGTCTATGCT CTACGTCATT TTGATGAGGG TAATTGTGAT ACATTAAG  
 13801 AAATACTCGT CACATACAAT TGCTGCTGATG ATGATTATT CAATAAGAAG  
 13851 GATTGGTATG ACTTCGAGA GAATCCTGAC ATCTTACCGG TATATGCTAA  
 13901 CTTAGGTGAG CGTGTACGCC AATCATTATT AAAGACTGTA CAATTCTGCG  
 13951 ATGCTATGCG TGATGCAAGG ATTGTAGGCG TACTGACATT AGATAATCAG  
 14001 GATCTTAATG GGAACGGTA CGATTTCGGT GATTCGTAC AAGTAGCACC  
 14051 AGGCTCGGGA GTTCCATTG TGATTTCATA TTACTCATG CTGATGCCA  
 14101 TCCTCACTT GACTAGGGCA TTGGCTGCTG AGTCCCATT GGATGCTGAT  
 14151 CTCGCAAAAC CACTTATTAA GTGGGATTG CTGAAATATG ATTTCACGGA  
 14201 AGAGAGACTT TGTCTCTCG ACCGTTATTAA TAAATATTGG GACCAGACAT  
 14251 ACCATCCCAA TTGTATTAAAC TGTTGGATG ATAGGTGAT CTTTCAATTG  
 14301 GCAAACTTTA ATGTGTTATT TTCTACTGTTG TTCCACCTA CAAGTTTGG  
 14351 ACCACTAGTA AGAAAAAATAT TTGTAGATGG TGTTCTTTT GTTGTTC  
 14401 CTGGATACCA TTTCTGTGAG TTAGGAGTCG TACATAATCA GGATGTAAC  
 14451 TTACATAGCT CGCGTCTCAG TTTCAAGGAA CTTTGTGATG ATGCTGCTGA  
 14501 TCCAGCTATG CATGCAGCTT CTGGCAATT ATTGCTAGAT AAACGCACTA  
 14551 CATGCTTTTC AGTAGCTGCA CTAACAAACA ATGTTGCTTT TCAAACGT  
 14601 AAACCCGGTA ATTAAATAAA AGACTTTAT GACTTTGCTG TGTCTAAAGG  
 14651 TTTCTTAAG GAAGGAAGTT CTGTTGAAC AAAACACTTC TTCTTGCTC  
 14701 AGGATGGCAA CGCTGCTATC AGTGATTATG ACTATTATCG TTATAATCTG  
 14751 CCAACAATGT GTGATATCAG ACAACTCTA TTCTGAGTTG AAGTTGTTGA  
 14801 TAAATACTTT GATTGTTACG ATGGTGGCTG TATTAAATGCC AACCAAGTAA  
 14851 TCGTTAACAA TCTGGATAAA TCAGCTGGTT TCCCATTAA TAAATGGG  
 14901 AAGGCTAGAC TTATTATGA CTCATGAGT TATGAGGATC AAGATGCACT  
 14951 TTTCGCGTAT ACTAAGCGTA ATGTCATCCC TACTATAACT CAAATGAATC  
 15001 TTAAGTATGC CATTAGTGCA AAGAATAGAG CTCGCACCGT AGCTGGTGT  
 15051 TCTATCTGTA GTACTATGAC AAATAGACAG TTTCATCAGA AATTATTGAA  
 15101 GTCAATAGCC GCCACTAGAG GAGCTACTGT GGTAAATTGGA ACAAGCAAGT  
 15151 TTTACGGTGG CTGGCATAAT ATGTTAAAAA CTGTTTACAG TGATGTTAGAA  
 15201 ACTCCACACC TTATGGGTTG GGATTATCCA AAATGTGACA GAGCCATGCC  
 15251 TAACATGCTT AGGATAATGG CCTCTCTGT TCTGCTCGC AAACATAACA  
 15301 CTGCTGTAA CTTATCACAC CGTTCTACA GGTTAGCTAA CGAGTGTGCG  
 15351 CAAGTATTAA GTGAGATGGT CATGTGTGGC GGCTCACTAT ATGTTAAACC  
 15401 AGGTGGAACA TCATCCGGTG ATGCTACAC TGCTTATGCT AATAGTGTCT  
 15451 TTAACATTG TCAAGCTGTT ACAGCCAATG TAAATGCACT TCTTCAACT  
 15501 GATGGTAATA AGATAGCTGA CAAGTATGTC CGCAATCTAC AACACAGGCT  
 15551 CTATGAGTGT CTCTATAGAA ATAGGGATGT TGATCATGAA TTGCTGGATG  
 15601 AGTTTACGC TTACCTGCGT AAACATTCT CCATGATGAT TCTTCTGAT  
 15651 GATGCCGTG TGTGCTATAA CAGTAACATG CGGGCTCAAG GTTGTAGTAGC  
 15701 TAGCATTAAAG AACTTTAAGG CAGTCTTTA TTATCAAAAT AATGTGTTCA

15751 TGTCTGAGGC AAAATGTTGG ACTGAGACTG ACCTTACTAA AGGACCTCAC  
15801 GAATTTGCT CACAGCATAAC AATGCTAGTT AAACAAGGGAG ATGATTACGT  
15851 GTACCTGCCT TACCCAGATC CATCAAGAAAT ATTAGGCCA GGCTGTTTG  
15901 TCGATGATAT TGTCAAAACA GATGGTACAC TTATGATTGA AAGGTTCTGT  
15951 TCACTGGCTA TTGATGCTTA CCCACTTACA AAACATCCTA ATCAGGAGTA  
16001 TGCTGATGTC TTTCACCTGT ATTACAATA CATTAGAAAG TTACATGATG  
16051 AGCTTACTGG CCACATGTTG GACATGTATT CCGTAATGCT AACTAATGAT  
16101 AACACCTCAC GGTACTGGGA ACCTGAGTTT TATGAGGCTA TGTACACACC  
16151 ACATACAGTC TTGCAGGCTG TAGGTGCTG TGTATTGTC AATTACAGA  
16201 CTTCACCTCG TTGCGGTGCC TGTATTAGGA GACCATTCCCT ATGTTGCAAG  
16251 TGCTGCTATG ACCATGTCTT CAACATCA CACAAATTAG TGTTGCTGT  
16301 TAATCCCTAT GTTGCAATG CCCCAGGTTG TGATGTCACT GATGTGACAC  
16351 AACTGTATCT AGGAGGTATG AGCTTATT GCAAGTCACA TAAGCCTCCC  
16401 ATTAGTTTC CATTATGTC TAATGTCAG GTTTTTGGTT TATACAAAAAA  
16451 CACATGTGTA GGCAGTGCAC ATGTCACTGA CTTCAATGCG ATAGCAACAT  
16501 GTGATGGAC TAATGCTGGC GATTACATAC TTGCCAACAC TTGTAUTGAG  
16551 AGACTCAAGC TTTT'CGCAGC AGAAAAGCCTC AAAGCCACTG AGGAAACATT  
16601 TAAGCTGTCA TATGGTATTG CTACTGTCAG CGAAGTACTC TCTGACAGAG  
16651 AATTGCATCT TTCAATGGGAG GTTGGAAAAC CTAGACCACC ATTGAACAGA  
16701 AACTATGTC TTACTGGTTA CGGTGTAAC AAAATAGTA AAGTACAGAT  
16751 TGGAGAGTAC ACCTTTGAAA AAGGTACTA TGGTGTGTC GTTGTGTACA  
16801 GAGGTACTAC GACATACAAG TTGAATGTTG GTGATTACCT TGTGTTGACA  
16851 TCTCACACTG TAATGCCACT TAGTGCACCT ACTCTAGTGC CACAAGAGCA  
16901 CTATGTGAGA ATTACTGGCT TGTACCCAAC ACTCAACATC TCAGATGAGT  
16951 TTTCTAGCAA TGTGCAAAT TATCAAAAGG TCGGCATGCA AAAGTACTCT  
17001 ACACCTCAAG GACCACCTGG TACTGGTAAG AGTCATTTG CCATCGGACT  
17051 TGCTCTCTAT TACCCATCTG CTCGCATAGT GTATACGGCA TGCTCTCATG  
17101 CAGCTGTTGA TGCCCTATGT GAAAAGGCAT TAAAATATTG GCCCATAGAT  
17151 AAATGTAGTA GAATCATACC TGCGCGTGC CGCGTAGAGT GTTTGATAA  
17201 ATTCAAAGTG AATTCAACAC TAGAACAGTA TGTTTCTGC ACTGTAAATG  
17251 CATTGCCAGA AACAACTGCT GACATTGAG TCTTGTGATGA AATCTCTATG  
17301 GCTACTAATT ATGACTTGAG TGTGTCAT GCTAGACTTC GTGAAAACA  
17351 CTACGTCTAT ATTGGCGATC CTGCTCAATT ACCAGCCCC CGCACATTGC  
17401 TGACTAAAGG CACACTAGAA CCAGAATATT TTAATTCACT GTGCAGACTT  
17451 ATGAAAACAA TAGGTCCAGA CATGTTCCCTT GGAACCTGTC GCCGTTGTCC  
17501 TGCTGAAATT GTTGACACTG TGAGTGCTTT AGTTTATGAC AATAAGCTAA  
17551 AAGCACACAA GGATAAGTCA GCTCAATGCT TCAAAATGTT CTACAAAGGT  
17601 GTTATTACAC ATGATGTTTC ATCTGCAATC AACAGACCTC AAATAGGCCT  
17651 TGTAAGAGAA TTCTTACAC GCAATCCTGC TTGGAGAAA GCTGTTTTA  
17701 TCTCACCTTA TAATTACAG AACGCTGTAG CTTCAAAAT CTTAGGATTG  
17751 CCTACGCAGA CTGTTGATTG ATCACAGGGT TCTGAATATG ACTATGTCAT  
17801 ATTACACACAA ACTACTGAAA CAGCACACTC TTGTAATGTC AACCGCTTCA  
17851 ATGTGCTAT CACAAGGGCA AAAATGGCA TTTTGTGCAAT AATGTCTGAT  
17901 AGAGATCTTT ATGACAAACT GCAATTACAA AGTCTAGAAA TACCACGTG  
17951 CAATGTGGCT ACATTACAAG CAGAAAATGT AACTGGACTT TTTAAGGACT  
18001 GTAGTAAGAT CATTACTGGT CTTCATCCTA CACAGGCACC TACACACCTC  
18051 AGCGTTGATA TAAAGTTCAA GACTGAAGGA TTATGTTGAC ATACACCAGG  
18101 CATAACAAAG GACATGACCT ACCGTAGACT CATCTCTATG ATGGGTTTCA  
18151 AAATGAATTA CCAAGTCAT GGTTACCTA ATATGTTTAT CACCCCGCGAA  
18201 GAAGCTTATC GTCACGTTCG TGCCTGGATT GGCTTGTGATG TAGAGGGCTG  
18251 TCATGCAACT AGAGATGCTG TGGGTACTAA CCTACCTCTC CAGCTAGGAT  
18301 TTTCTACAGG TGTAACTTA GTAGCTGTAC CGACTGGTTA TGTGACACT  
18351 GAAAATAACA CAGAATTACAC CAGAGTTAAT GCAAAACCTC CACCAAGGTGA

18401 CCAGTTAAA CATCTTATAC CACTCATGTA TAAAGGCTTG CCCTGGAATG  
 18451 TAGTGCAT TAAGATAGTA CAAATGCTCA GTGATACACT GAAAGGATTG  
 18501 TCAGACAGAG TCGTGTTCGT CCTTTGGCG CATGGCTTG AGCTTACATC  
 18551 AATGAAGTAC TTTGTCAGA TTGGACCTGA AAGAACGTGT TGTCTGTG  
 18601 ACAAACGTGC AACTTGCTT TCTACTTCAT CAGATACTTA TGCCCTGCTGG  
 18651 AATCATTCTG TGGGTTTGA CTATGTCTAT AACCCATTAA TGATTGATGT  
 18701 TCAGCAGTGG GGCTTACCG GTAACCTTCAGA GAGTAACCAT GACCAACATT  
 18751 GCCAGGTACA TGGAAATGCA CATGTGGCTA GTTGTGATGC TATCATGACT  
 18801 AGATGTTAG CAGTCCATGA GTGCTTGTT AAGCGCGTTG ATTGGTCTGT  
 18851 TGAATACCCCT ATTATAGGAG ATGAACTGAG GTTAAATTCT GCTTGCAGAA  
 18901 AAGTACAACA CATGGTTGTG AAGTCTGCAT TGCTTGCTGA TAAGTTCCA  
 18951 GTTCTTCATG ACATGGAAA TCCAAAGGCT ATCAAGTGTG TGCCTCAGGC  
 19001 TGAAGTAGAA TGGAAAGTTCT ACGATGCTCA GCCATGTAGT GACAAAGCTT  
 19051 ACAAAATAGA GGAGCTCTTC TATTCTTATG CTACACATCA CGATAAAATTC  
 19101 ACTGATGGTG TTTGTTGTT TTGGATTGT AACGTTGATC GTTACCCAGC  
 19151 CAATGCAATT GTGTGTAGGT TTGACACAAG AGCCTTGTC AACTTGAAC  
 19201 TACCAAGGCTG TGATGGTGGT AGTTTGATG TGAATAAGCA TGCATTCCAC  
 19251 ACTCCAGCTT TCGATAAAAG TGCATTTACT AATTAAAGC AATTGCCCTT  
 19301 CTTTTACTAT TCTGATAGTC CTTGTGAGTC TCATGGAAA CAAGTAGTGT  
 19351 CGGATATTGA TTATGTTCCA CTCAAATCTG CTACGTGTAT TACACGATGC  
 19401 AATTAGGTG GTGCTGTTG CAGACACCAT GCAAATGAGT ACCGACAGTA  
 19451 CTTGGATGCA TATAATATGA TGATTTCTGC TTGGATTAGC CTATGGATT  
 19501 ACAAAACAATT TGATACTTAA AACCTGTGGA ATACATTAC CAGGTTACAG  
 19551 AGTTAGAAA ATGTGGCTTA TAATGTTGTT AATAAAGGAC ACTTTGATGG  
 19601 ACACGCCGGC GAAGCACCTG TTTCCATCAT TAATAATGCT GTTTACACAA  
 19651 AGGTAGATGG TATTGATGTG GAGATCTTG AAAATAAGAC AACACTTCCT  
 19701 GTTAATGTTG CATTGAGCT TTGGGCTAAG CGTAACATTA AACCAGTGCC  
 19751 AGAGATTAAG ATACTCAATA ATTTGGGTGT TGATATGCT GCTAACTG  
 19801 TAATCTGGGA CTACAAAAGA GAAGCCCCAG CACATGTATC TACAATAGGT  
 19851 GTCTGCACAA TGACTGACAT TGCCAAGAAA CCTACTGAGA GTGCTTGTTC  
 19901 TTCACTTACT GTCTGTTTG ATGGTAGAGT GGAAGGACAG GTAGACCTTT  
 19951 TTAGAAACGC CGGTAAATGGT GTTTAATAA CAGAAGGTTG AGTCAAAGGT  
 20001 CTAAACACCTT CAAAGGGACC AGCACAAGCT AGCGTCATG GAGTCACATT  
 20051 AATTGGAGAA TCAGTAAAAA CACAGTTAA CTACTTTAAG AAAGTAGACG  
 20101 GCATTATTCA ACAGTTGCCT GAAACACTTAA TACTCAGAG CAGAGACTTA  
 20151 GAGGATTTA AGCCCAGATC ACAAATGGAA ACTGACTTTC TCGAGCTCGC  
 20201 TATGGATGAA TTCAACAGC GATATAAGCT CGAGGGCTAT GCCTCGAAC  
 20251 ACATCGTTA TGGAGATTTC AGTCATGGAC ACTTGGCGG TCTTCATTAA  
 20301 ATGATAGGCT TAGCCAAGCG CTCACAAGAT TCACCACTTA AATTAGAGGA  
 20351 TTTTATCCCT ATGGACAGCA CAGTAAAAAA TTACTTCATA ACAGATGCGC  
 20401 AAACAGGTTCACTAAATGT GTGTGTTCTG TGATTGATCT TTTACTTGAT  
 20451 GACTTTGTCG AGATAATAAA GTCACAAGAT TTGTCAGTGA TTTCAAAGT  
 20501 GGTCAAGGTT ACAATTGACT ATGCTGAAAT TTCAATTGATC TTTGGGTGA  
 20551 AGGATGGACA TGTTGAAACC TTCTACCCAA AACTACAAGC AAGTCAAGCG  
 20601 TGGCAACCCAG GTGTTGCGAT GCCTAACTTG TACAAGATGC AAAGAATGCT  
 20651 TCTTGAAAAG TGTGACCTTC AGAATTATGG TGAAAATGCT GTTATACCAA  
 20701 AAGGAATAAT GATGAATGTC GCAAAGTATA CTCAACTGTG TCAATACCTTA  
 20751 AATACACTTA CTTTAGCTGT ACCCTACAAC ATGAGAGTTA TTCACTTGG  
 20801 TGCTGGCTCT GATAAAGGAG TTGCAACCAAG TACAGCTGTG CTCAGACAAAT  
 20851 GGTTGCCAAC TGGCACACTA CTTGTCGATT CAGATCTAA TGACTTCGTC  
 20901 TCCGACGCGAG ATTCTACTTT AATTGGAGAC TGTGCAACAG TACATACGGC  
 20951 TAATAAATGG GACCTTATTA TTAGCGATAT GTATGACCTT AGGACCAAAC  
 21001 ATGTGACAAA AGAGAATGAC TCTAAAGAAG GGTTTTCAC TTATCTGTG

21051 GGATTTATAA AGCAAAAAGT AGCCCTGGGT GGTTCTATAG CTGTAAGAT  
21101 AACAGAGCAT TCTTGGAAATG CTGACCTTTA CAAGCTTATG GGCCATTCT  
21151 CATGGTGGAC AGCTTTGTT ACAAAATGTA ATGCATCATC ATCGGAAGCA  
21201 TTTTTAATTG GGGCTAACTA TCTTGGCAAG CCGAAGGAAC AAATTGATGG  
21251 CTATACCATG CATGCTAACT ACATTTCTG GAGGAACACA AATCCATATCC  
21301 AGTTGTCTTC CTATTCACTC TTTGACATGA GCAAATTCC TCTTAAATTA  
21351 AGAGGAACGTG CTGTAATGTC TCTTAAGGAG AATCAAATCA ATGATATGAT  
21401 TTATTCTCTT CTGGAAAAAG GTAGGCTTAT CATTAGAGAA ACAACAGAG  
21451 TTGTGGTTTC AAGTGTATT CTTGTTAACAA ACTAAACGAA CATGTTTATT  
21501 TTCTTATTAT TTCTTACTCT CACTAGGGT AGTGACCTTG ACCGGTGCAC  
21551 CACTTTGAT GATGTTCAAG CTCCTAATTA CACTCAACAT ACTTCATCTA  
21601 TGAGGGGGGT TTACTATCCT GATGAAATT TTAGATCAGA CACTTTAT  
21651 TTAACTCAGG ATTTATTTCT TCCATTAT TCTAATGTTA CAGGGTTCA  
21701 TACTATTAAT CATACTTTG GCAACCCGT CATACTTT AAGGATGGTA  
21751 TTTATTTGC TGCCACAGAG AAATCAAATG TTGTCGTGG TTGGGTTTT  
21801 GGTTCTACCA TGAACAAACAA GTCACAGTCG GTGATTATTA TTAACAATT  
21851 TACTAATGTT GTTATACGAG CATGTAACCT TGAATTGTTG GACAACCCCT  
21901 TCTTGCTGT TTCTAAACCC ATGGGTACAC AGACACATAC TATGATATT  
21951 GATAATGCAT TTAATTGCAC TTTGAGTAC ATATCTGATG CCTTTCGCT  
22001 TGATGTTCA GAAAAGTCAG GTAATTAA ACACCTACGA GAGTTGTGT  
22051 TTAAAAATAA AGATGGGTT CTCTATGTTT ATAAGGGCTA TCAACCTATA  
22101 GATGTAGTTG GTGATCTACC TTCTGGTTT AACACTTGA AACCTATTT  
22151 TAAGTTGCCT CTTGGTATTAA ACATTACAAA TTTAGAGCC ATTCTTACAG  
22201 CCTTTTCACC TGCTCAAGAC ATTTGGGCA CGTCAGCTGC AGCCTATTT  
22251 GTTGGCTATT TAAAGCCAAC TACATTATG CTCAAGTATG ATGAAAATGG  
22301 TACAATCACA GATGCTGTTG ATTGTTCTCA AAATCCACTT GCTGAACTCA  
22351 AATGCTCTGT TAAGAGCTT GAGATTGACA AAGGAATTAA CCAGACCTCT  
22401 AATTTCAGGG TTGTTCCCTC AGGAGATGTT GTGAGATTCC CTAATATTAC  
22451 AAACTTGTGT CTTTGGAG AGGTTTTAA TGCTACTAAA TTCCCTCTG  
22501 TCTATGCTATG GGAGAGAAAA AAAATTCTA ATTGTGTTGC TGATTACTCT  
22551 GTGCTCTACA ACTCAACATT TTTCAACC TTAAGTGT ATGGCGTTTC  
22601 TGCCACTAAG TTGAATGATC TTTGCTTCTC CAATGTCTAT GCAGATTCTT  
22651 TTGTAGTCAGA GGGAGATGAT GTAAGACAAA TAGGCCAGG ACAAACTGGT  
22701 GTTATTGCTG ATTATAATTA TAAATTGCCA GATGATTCA TGGGTTGTGT  
22751 CCTTGCTTGG AATACTAGGA ACATTGATGC TACTCAACT GGTATTATA  
22801 ATTATAAATA TAGGTATCTT AGACATGGCA AGCTTAGGCC CTTGAGAGA  
22851 GACATATCTA ATGTGCCCTT CTCCCCGTAT GGCAAACCTT GCACCCACC  
22901 TGCTCTTAAT TGTTATTGGC CATTAAATGA TTATGGTTT TACACCACTA  
22951 CTGGCATTGG CTACCAACCT TACAGAGTTG TAGTACTTTC TTTGAACCT  
23001 TAAATGCAC CGGCCACGGT TTGTGGACCA AAATTATCCA CTGACCTTAT  
23051 TAAGAACCAAG TGTGTCAATT TTAATTAA TGGACTCACT GGTACTGGT  
23101 TGTTAACTCC TTCTTCAAAG AGATTCAAC CATTCAACA ATTTGCCGT  
23151 GATGTTCTG ATTCACTGA TTCCGTTCGA GATCCTAAA CATCTGAAAT  
23201 ATTAGACATT TCACCTTGTCTTGGGGGG TGTAAGTGT ATTACACCTG  
23251 GAACAAATGC TTCACTGAA GTTGTGTTC TATATCAAGA TGTTAACTGC  
23301 ACTGATGTTT CTACAGCAAT TCATGCGAGT CAACTCACAC CAGCTGGCG  
23351 CATATATTCT ACTGGAAACA ATGTATTCCA GACTCAAGCA GGCTGTCTTA  
23401 TAGGAGCTGA GCATGTCGAC ACTTCTTATG AGTGCAGACAT TCCTATTGGA  
23451 GCTGGCATTG TGCTAGTTA CCATACAGTT TCTTATTAC GTAGTACTAG  
23501 CAAAAAAATCT ATTGTGGCTT ATACTATGTC TTAGGTGCT GATAGTTCAA  
23551 TTGCTTACTC TAATAACACC ATTGCTATAC CTACTAACTT TTCAATTAGC  
23601 ATTACTACAG AAGTAATGCC TGTTCTATG GCTAAAACCT CCGTAGATTG  
23651 TAATATGTAC ATCTGCGGAG ATTCTACTGA ATGTGCTAAT TTGCTTCTCC

23701 AATATGGTAG CTTTGACA CAACTAAATC GTGCACTCTC AGGTATTGCT  
 23751 GCTAACAGG ATCGAACAC ACGTGAAGTG TTCGCTCAAG TCAAACAAAT  
 23801 GTACAAAACC CCAACTTGA AATATTTGG TGGTTTAAT TTTTACAAA  
 23851 TATTACCTGA CCCTCTAAAG CCAACTAAGA GGTCTTTAT TGAGGACTTG  
 23901 CTCTTTAATA AGGTGACACT CGCTGATGCT GGCTCATGA AGCAATATGG  
 23951 CGAATGCCA GGTGATATTA ATGCTAGAGA TCTCATTGT GCGCAGAAGT  
 24001 TCAATGGACT TACAGTGTG CCACCTCTGC TCACTGATGA TATGATTGCT  
 24051 GCCTACACTG CTGCTCTAGT TAGTGGTACT GCCACTGCTG GATGGACATT  
 24101 TGGTGCCTGGC GCTGCTCTC AAATACCTT TGCTATGCAA ATGGCATATA  
 24151 GGTTCAATGG CATTGGAGTT ACCCAAAATG TTCTCTATGA GAACCAAAAA  
 24201 CAAATGCCA ACCAATTAA CAAGGCGATT AGTCAAATTCAAGAATCACT  
 24251 TACAACAACA TCAACTGCAT TGGGCAAGCT GCAAGACGTT GTTAACCAGA  
 24301 ATGCTCAAGC ATAAACACA CTTGTTAAC AACTTAGCTC TAATTTGGT  
 24351 GCAATTCAA GTGTGCTAAA TGATATCCTT TCGCGACTTG ATAAAGTCGA  
 24401 GGCAGGGTA CAAATTGACA GGTTAATTAC AGGCAGACTT CAAAGCCTTC  
 24451 AAACCTATGT AACACAACAA CTAATCAGGG CTGCTGAAAT CAGGGCTTCT  
 24501 GCTAATCTG CTGCTACTAA AATGCTGAG TGTGTTCTG GACAATCAA  
 24551 AAGAGTTGAC TTTGTGGAA AGGGCTACCA CTTATGTC TTCCCACAAG  
 24601 CAGCCCCGCA TGGTGTGTC TTCTACATG TCACGTATGT GCCATCCCAG  
 24651 GAGAGGAAC TCAACCACAGC GCCAGCAATT TGTGATGAAAG GCAAAGCATA  
 24701 CTTCCCTCGT GAAGGGTTT TTGTTAA TGGCACTCT TGTTTATTA  
 24751 CACAGAGGAA CTTCTTTCT CCACAAATAA TTACTACAGA CAATACATT  
 24801 GTCTCAGGAA ATTGTGATGT CGTTATTGGC ATCATTAAACA ACACAGTTA  
 24851 TGATCCTCTG CAACCTGAGC TCGACTCATT CAAAGAAGAG CTGGACAAGT  
 24901 ACTTCAAAAA TCATACATCA CCAGATGTTG ATCTTGGCGA CATTTCAGGC  
 24951 ATTAACGCTT CTGCGTCAA CATTCAAAA GAAATTGACC GCCTCAATGA  
 25001 GGTCGCTAAA AATTTAAATG AATCACTCAT TGACCTTCAA GAATTGGAA  
 25051 AATATGAGCA ATATATTAAA TGGCCTTGGT ATGTTGGCT CGGCTTCATT  
 25101 GCTGGACTAA TTGCCATCGT CATGGTTACA ATCTTGCTT GTTGCATGAC  
 25151 TAGTTGTTGC AGTTGCTCA AGGGTGCATG CTCTTGTGGT TCTTGCTGCA  
 25201 AGTTTGATGA GGATGACTCT GAGCCAGTTC TCAAGGGTGT CAAATTACAT  
 25251 TACACATAAA CGAACTTATG GATTTGTTA TGAGATTTTT TACTCTGGA  
 25301 TCAATTACTG CACAGCCAGT AAAAATTGAC AATGCTTCTC CTGCAAGTAC  
 25351 TGTCATGCT ACAGCAACGA TACCGCTACA AGCCTCACTC CCTTCGGAT  
 25401 GGCTTGTAT TGGCGTTGCA TTTCTTGTG TTTTCAGAG CGCTACAA  
 25451 ATAATTGCGC TCAATAAAAG ATGGCAGCTA GCCCTTATA AGGGCTTCCA  
 25501 GTTCATTTGC AATTTACTGC TGCTATTGT TACCATCTAT TCACATCTT  
 25551 TGCTTGTGCG TGCAGGTATG GAGGGCGAAT TTTGTACCT CTATGCCTG  
 25601 ATATATTTTC TACAATGCAT CAACGCATGT AGAATTATTA TGAGATGTTG  
 25651 GCTTTGTTGG AAGTGCAAAT CCAAGAACCC ATTACTTTAT GATGCCAACT  
 25701 ACTTTGTTTG CTGGCACACA CATAACTATG ACTACTGTAT ACCATATAAC  
 25751 AGTGTACAG ATACAATTGT CGTTACTGAA GGTGACGGCA TTTCAACACC  
 25801 AAAACTCAA GAAGACTACC AAATTGGTGG TTATTCTGAG GATAGGCAC  
 25851 CAGGTGTTAA AGACTATGTC GTTGACATG GCTATTGAC CGAAGTTAC  
 25901 TACAGCTTG AGTCTACACA AATTACTACA GACACTGGTA TTGAAAATGC  
 25951 TACATTCTC ATCTTAAACA AGCTTGTAA AGACCCACCG AATGTGCAA  
 26001 TACACACAAT CGACGGCTCT TCAGGAGTTG CTAATCCAGC AATGGATCCA  
 26051 ATTTATGATG AGCCGACGAC GACTACTAGC GTGCCTTGT AAGCACAAGA  
 26101 AAGTGAAGTAC GAACTTATGT ACTCATTGCT TTGGAAAGAA ACAGGTACGT  
 26151 TAATAGTTAA TAGCGTACTT CTTTTCTTG CTTCGTGGT ATTCTTGCTA  
 26201 GTCACACTAG CCATCCTTAC TGCGCTTCGA TTGTTGCGT ACTGCTGCAA  
 26251 TATTGTTAAC GTGAGTTAG TAAAACCAAC GGTTTACGTC TACTCGCGTG  
 26301 TTAACATCT GAACTCTTCT GAAGGGAGTTC CTGATCTTCT GGTCTAAACG

26351 AACTAACTAT TATTATTATT CTGTTGGAA CTAAACATT GCTTATCATG  
26401 GCAGACAACG GTACTATTAC CGTTGAGGAG CTTAAACAAC TCCTGGAACA  
26451 ATGGAACCTA GTAATAGGTT TCCTATTCT AGCCTGGATT ATGTTACTAC  
26501 AATTGCGCTA TTCTAATCGG AACAGGTTT TGTACATAAT AAAGCTTGT  
26551 TTCTCTGGC TCTTGTGGC AGTAACACTT GCTTGTGTTG TGCTTGTG  
26601 TGTCTACAGA ATTAATTGGG TGACTGGCGG GATTGCGATT GCAATGGCTT  
26651 GTATTGTAGG CTTGATGTGG CTTAGCTACT TCGTTGCTTC CTTCAGGCTG  
26701 TTTGCTCGTA CCCGCTCAAT GTGGTCATTCAACCCAGAAA CAAACATTCT  
26751 TCTCAATGTG CCTCTCCGGG GGACAAATTGT GACCAGACCG CTCATGGAAA  
26801 GTGAACTTGT CATTGGTGT GTGATCATTG TGTTGCACTT GCGAATGGCC  
26851 GGACACCCCCC TAGGGCGCTG TGACATTAAG GACCTGCCAA AAGAGATCAC  
26901 TGTGGCTACA TCACGAACGC TTTCTTATTA CAAATTAGGA GCCTCGCAGC  
26951 GTGTAGGCAC TGATTCAAGT TTTGCTGCAT ACAACCGCTA CCGTATTGGA  
27001 AACTATAAAT TAAATACAGA CCACGCCGGT AGCAACGACA ATATTGCTT  
27051 GCTAGTACAG TAAGTGCACAA CAGATGTTTC ATCTTGTGA CTTCCAGGTT  
27101 ACAATAGCAG AGATATTGTAT TATCATTATG AGGACTTTCA GGATTGCTAT  
27151 TTGGAATCTT GACGTTATAA TAAGTTCAAT AGTGAGACAA TTATTAAAGC  
27201 CTCTAACTAA GAAGAATTAT TC GGAGTTAG ATGATGAGA ACCTATGGAG  
27251 TTAGATTATC CATAAAACGA ACATGAAAAT TATTCTCTTC CTGACATTGA  
27301 TTGTATTTCAT ATCTTGCAG CTATATCACT ATCAGGAGTG TGTTAGAGGT  
27351 ACGACTGTAC TACTAAAAGA ACCTTGCCTA TCAGGAACAT ACGAGGGCAA  
27401 TTCACCATTT CACCCCTCTTG CTGACAAATAA ATTGCACTA ACTTGCCTA  
27451 GCACACACTT TGCTTTGCT TGTGCTGACG GTACTCGACA TACCTATCAG  
27501 CTGCGTCAA GATCAGTTTC ACCAAAACCTT TTCACTCAGAC AAGAGGAGGT  
27551 TCAACAAAGAG CTCTACTCGC CACTTTCTT CATTGTTGCT GCTCTAGTAT  
27601 TTTTAATACT TTGCTTCACC ATTAAGAGAA AGACAGAATG AATGAGCTCA  
27651 CTTTAATTGA CTTCATTTTG TGCTTTTAG CCTTTCTGCT ATTCTTGTT  
27701 TTAATAATGC TTATTATATT TTGGTTTCA CTCGAAATCC AGGATCTAGA  
27751 AGAACCTTGT ACCAAAGTCT AAACGAACAT GAAACTTCTC ATTGTTTGA  
27801 CTTGTATTC TCTATGCAGT TGCAATATGCA CTGTTAGTACA GCGCTGTGCA  
27851 TCTAATAAAC CTCATGTGCT TGAAGATCCT TGTAAGGTAC AACACTAGGG  
27901 GTAATACTTA TAGCACTGCT TGGCTTGTG CTCTAGGAAA GGTTTACCT  
27951 TTTCATAGAT GGCACACTAT GGTTCAAACA TGCAACACCTA ATGTTACTAT  
28001 CAACTGTCAA GATCCAGCTG GTGGTGCCT TATAGCTAGG TGTTGGTACC  
28051 TTCATGAAGG TCACCAAAC GCTGCATTAA GAGACGTACT TGTTGTTTA  
28101 AATAAACGAA CAAATTAAAA TGTCTGATAA TGGACCCCAA TCAAACCAAC  
28151 GTAGTGCCTT CCGCATTACA TTTGGTGGAC CCACAGATTC AACTGACAAT  
28201 AACCGAAATG GAGGACGCAA TGGGGCAAGG CCAAACAGC GCCGACCCCA  
28251 AGGTTTACCC ATAATACTG CGTCTGGTT CACAGCTCTC ACTCAGCATG  
28301 GCAAGGAGGA ACTTAGATTC CCTCGAGGCC AGGGCGTTCC AATCAACACC  
28351 AATAGTGGTC CAGATGACCA AATTGGCTAC TACCGAAGAG CTACCCGACG  
28401 AGTTGGTGGT GGTGACGGCA AAATGAAAGA GCTCAGCCCC AGATGGTACT  
28451 TCTATTACCT AGGAACATGGC CCAGAAGCTT CACTTCCCTA CGGCGCTAAC  
28501 AAAGAAGGCA TCGTATGGGT TGCAACTGAG GGAGCCTTGA ATACACCCAA  
28551 AGACCACATT GGCACCCGCA ATCCTAATAA CAATGCTGCC ACCGTGCTAC  
28601 AACTTCTCA AGGAACAAACA TTGCCAAAAG GCTTCTACGC AGAGGGAAAGC  
28651 AGAGGGGGCA GTCAAGCCTC TTCTCGCTCC TCATCACGTA GTCGCGGTAA  
28701 TTCAAGAAAT TCAACTCCTG GCAGCAGTAG GGGAAATTCT CCTGCTCGAA  
28751 TGGCTAGCGG AGGTGGTGA ACTGCCCTCG CGCTATTGCT GCTAGACAGA  
28801 TTGAACCAGC TTGAGAGCAA AGTTCTGGT AAAGGCCAAC AACACACAAGG  
28851 CCAAACGTCA ACTAAGAAAT CTGCTGCTGA GGCATCTAAA AAGCTCGCC  
28901 AAAAACGTAC TGCCACAAAA CAGTACAACG TCACTCAAGC ATTTGGGAGA  
28951 CGTGGTCCAG AACAAACCCA AGGAAATTG GGGGACCAAG ACCTAATCAG

29001 ACAAGGAAC T GATTACAAAC ATTGGCCGCA AATTGCACAA TTTGCTCCAA  
29051 GTGCCTCTGC ATTCTTTGGA ATGTCACGCA TTGGCATGGA AGTCACACCT  
29101 TCGGGAACAT GGCTGACTTA TCATGGAGCC ATAAATTGG ATGACAAAGA  
29151 TCCACAAATT AAAGACAACG TCATACTGCT GAACAAGCAC ATTGACGCAT  
29201 ACAAAACATT CCCACCAACA GAGCCTAAAA AGGACAAAAA GAAAAAGACT  
29251 GATGAAGCTC AGCCTTGCC GCAGAGACAA AAGAAGCAGC CCACTGTGAC  
29301 TCTTCTTCCT GCGGCTGACA TGGATGATT CTCCAGACAA CTTCAAAATT  
29351 CCATGAGTGG AGCTTCTGCT GATTCAACTC AGGCATAAAC ACTCATGATG  
29401 ACCACACAAG GCAGATGGGC TATGTAAACG TTTTCGCAAT TCCGTTACG  
29451 ATACATAGTC TACTCTTGTG CAGAATGAAT TCTCGTAAC AACACAGCACA  
29501 AGTAGGTTA GTTAACCTTA ATCTCACATA GCAATCTTTA ATCAATGTGT  
29551 AACATTAGGG AGGACTTGAA AGAGCCACCA CATTTCATC GAGGCCACGC  
29601 GGAGTACGAT CGAGGGTACA GTGAATAATG CTAGGGAGAG CTGCCTATAT  
29651 GGAAGAGCCC TAATGTGTA AATTAATTAGT AGTAGTGCTA TCCCCATGTG  
29701 ATTTTAATAG CTTCTTAGGA GAATGAC

Group 1 Coronavirus (Human Coronavirus-HCoV 229E and Procine epidemic diarrhea virus-PEDV) with SARS (3 different isolates: Toronto, Singapore and Guangzhou)

SEQ ID NO: 23

NC_004718 (TOR)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
AY283794 (GIS)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
AY278489 (G2)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
NC_002645 (HCoV)	AATATGTTAAAAGAACCTGATGCCGATGTGATGATCTAAATTGATGGGATGGGACTAT
NC_003436 (PEDV)	AATATGCTTAAGAACCTTATTGATGGTGTGAAAATCCGTGCTTATGGGTTGGGATTAC
NC_004718 (TOR)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
AY283794 (GIS)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
AY278489 (G2)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
NC_002645 (HCoV)	CCTAAGTGTGATAGAGCTATGCCCTCAATGATCTGTTATGGTGTGGGATGATCTTAGGT
NC_003436 (PEDV)	CCAAAATGTGCGATAGAGCACTGCCAACATGATACGCGATGATTAGCCATGATCTAGGC
NC_004718 (TOR)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
AY283794 (GIS)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
AY278489 (G2)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
NC_002645 (HCoV)	TCTAAGCATGTCACATGTTGACGGCTAGTGTGATAATTAGACTTAGTAAATGAGCTT
NC_003436 (PEDV)	TCTAAGCACACACATGCTGAGTTCTACTGACCGCTTTTCAAGGTTGTGCAATGAATTG

SEQ ID NO: 24

NC_004718 (TOR)	GGCGCAAGTATTAAGTGAGATGGTCATGTGTCGGCGCTCACTATATGTTAAACCAAGGTGGA
AY283794 (GIS)	GGCGCAAGTATTAAGTGAGATGGTCATGTGTCGGCGCTCACTATATGTTAAACCAAGGTGGA
AY278489 (G2)	GGCGCAAGTATTAAGTGAGATGGTCATGTGTCGGCGCTCACTATATGTTAAACCAAGGTGGA
NC_002645 (HCoV)	GCTCAAGTTTTGACCGAGCTGTTTATTCAATGGTGGGTTTTATTAAACCTGGTGTGTT
NC_003436 (PEDV)	GCTCAAGTCTTACTGAGGTGTTTATTCTAATGGAGGGTTTTATTGAAGCCAGGTGTT

Group 2 Coronaviruses (Bovine coronavirus-BCoV and Murine hepatitis virus-MHV) with SARS (3 isolates: Singapore, Guangzhou and Toronto)

SEQ ID NO: 23

AY283794 (GIS)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
AY278489 (G2)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
NC_004718 (TOR)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
NC_003045 (BCoV)	GATATGTTACGCGCTTATTAAAGATGTTGATAATCCCTGACTTATGGGTTGGGATTAT
NC_001846 (MHV)	GATATGTTACGCGCTTATTAAAGATGTTGATAATCCCTGACTTATGGGTTGGGACTAT
AY283794 (GIS)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
AY278489 (G2)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
NC_004718 (TOR)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
NC_003045 (BCoV)	CCTAAGTGTGATCGCTGCTAGCCTAACATACGTTAGCTTGTAGTAGCTGGTTTGGCT
NC_001846 (MHV)	CCTAAGTGTGATCGCTGCTAGCCTAACATACGCTTGTAGTAGCTGGCTAGCC
AY283794 (GIS)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
AY278489 (G2)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
NC_004718 (TOR)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
NC_003045 (BCoV)	CGCAAAACATGAGGCATGTTGCTGCAAAAGCGATAGGTTTATCGACTTGCAGAATGAATGC
NC_001846 (MHV)	CGTAAACATGATTGTCGCTGTTCCATACGGGATAGATTCTATCGTCTTGCAGAAGTC

SEQ ID NO: 24

AY283794 (GIS)	GGCGCAAGTATTAAGTGAGATGGTCATGTGTCGGCGCTCACTATATGTTAAACCAAGGTGGA
AY278489 (G2)	GGCGCAAGTATTAAGTGAGATGGTCATGTGTCGGCGCTCACTATATGTTAAACCAAGGTGGA
NC_004718 (TOR)	GGCGCAAGTATTAAGTGAGATGGTCATGTGTCGGCGCTCACTATATGTTAAACCAAGGTGGA
NC_003045 (BCoV)	GCACAAGTTCTGACTGAAATTGTTATGTGTCGGCTGTATTATGTTAAAGCCTGGTGGC
NC_001846 (MHV)	GCCCAGTTTGAGTGAATTGTTATGTGTCGGTGGTTGTATTATGTTAAACCAAGGTGCG

Group 3 Coronavirus (Avian Infectious Bronchitis Virus-IBV) with SARS (3 isolates: Singapore, Guangzhou and Toronto)

SEQ ID NO: 23

AY283794 (GIS)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
AY278489 (G2)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
NC_004718 (TOR)	AATATGTTAAAACGTGTTACAGTGATGTAGAAAACCTCCACACCTTATGGGTTGGGATTAT
NC_001451 (IBV)	AAACATGTTGAGAACCTGATTAGGGTGTGAGAGCCAAATTCTTATGGGTTGGGATTAT
AY283794 (GIS)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
AY278489 (G2)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
NC_004718 (TOR)	CCAAAATGTGACAGAGCCATGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
NC_001451 (IBV)	CCTAAGTGTGATAGAGCACTGCCAACATGCTTAGGATAATGGCCTCTTGTCTTGCT
AY283794 (GIS)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
AY278489 (G2)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
NC_004718 (TOR)	CGCAAAACATAACACTTGCTGTAACCTTATCACCCGTTCTACAGGTTAGCTAACGAGTGT
NC_001451 (IBV)	CGCAAAACACACTAACCTGTTAGTGTGCTGAGCGCTTATAGGTTGTATAATGAAATGC

AY283794 (GIS)  
AY278489 (G2)  
NC\_004718 (TOR)  
NC\_001451 (IBV)

SEQ ID NO: 24  
GGCGCAAGTATTAAGTGAGATGGTCATGTGTGGCGGCTCAC-TATATGTTAACCGGTGG  
GGCGCAAGTATTAAGTGAGATGGTCATGTGTGGCGGCTCAC-TATATGTTAACCGGTGG  
GGCGCAAGTATTAAGTGAGATGGTCATGTGTGGCGGCTCAC-TATATGTTAACCGGTGG  
GCCCGAGGTCTTATCTGAAACTGT-ACTTGCTACAGGTGGTATTTATGTTAACCTGGTGG

